STIC-Biotech/ChemLib 118457

From: Whiteman, Brian
Sent: Thursday, April 01, 2004 1:00 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

09/845,416 4/30/01 Xiao Xiao
4/30/01

search nucleotides 900-3000 of SEQ ID NO: 2 against us patent and us patent application databases

search nucleotides 1000-3000 of SEQ ID NO: 6 against us patent and us patent application databases

search nucleotides 600-3000 of SEQ ID NO: 9 against us patent and us patent application databases

search nucleotides 1800-3100 of SEQ ID NO: 10 against us patent and us patent application databases

search nucleotides 900-3000 of SEQ ID NO: 12 against us patent and us patent application databases

search nucleotides 900-2500 of SEQ ID NO: 14 against us patent and us patent application databases

Thanks,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

Searcher:_____	TYPE OF SEARCH:	VENDOR/COST (where applic.)
Phone:_____	NA Sequences:_____	STN:_____
Location:_____	AA Sequences:_____	DIALOG:_____
Date Picked Up:_____	Structures:_____	Questel/Orbit:_____
Date Completed:_____	Bibliographic:_____	DRLink:_____
Searcher Prep/Review:____	Litigation:_____	Lexis/Nexis:_____
Clerical:_____	Full text:_____	Sequence Sys.:_____
Online time:_____	Patent Family:_____	WWW/Internet:_____
	Other:_____	Other (specify):_____

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

un on:      April  4, 2004, 11:55:32 ; Search time 516.337 Seconds
                                        (without alignments)
                                        16463.377 Million cell updates/sec

itle:       US-09-845-416-6_COPY_1000_3000
erfect score: 2001
equence:    1 ggcagttcattgatggagag.......tcaaccacgagactcaaaca 2001

coring table: IDENTITY_NUC
            Gapop 10.0 , Gapext 1.0

earched:    3373863 seqs, 212409041 residues

otal number of hits satisfying chosen parameters:      6747726

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :   N_Geneseq_29Jan04:*
            1: geneseqn1980s:*
            2: geneseqn1990s:*
            3: geneseqn2000s:*
            4: geneseqn2001as:*
            5: geneseqn2001bs:*
            6: geneseqn2002s:*
            7: geneseqn2003as:*
            8: geneseqn2003bs:*
            9: geneseqn2003cs:*
           10: geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| | | % Query | | | | |
esult No.	Score	Match	Length	DB	ID	Description
1	2001	100.0	3999	6	AAD37234	Aad37234 Human dys
2	2001	100.0	4966	6	AAD37256	Aad37256 Adeno-ass
3	2001	100.0	4990	6	AAD37262	Aad37262 Adeno-ass
4	1709	85.4	3858	6	AAD37237	Aad37237 Human dys
5	1709	85.4	4825	6	AAD37257	Aad37257 Adeno-ass
6	1709	85.4	4848	6	AAD37263	Aad37263 Adeno-ass
7	1709	85.4	5060	6	AAD37264	Aad37264 Adeno-ass
8	1662.8	83.1	4182	6	AAD37230	Aad37230 Human dys
9	1662.8	83.1	5149	6	AAD37255	Aad37255 Adeno-ass
10	1283	64.1	5462	6	ABK81999	Abk81999 DNA encod
11	1182.6	59.1	5952	6	AAD06794	Aad06794 Human dys
12	1182.6	59.1	8689	6	ABK82000	Abk82000 DNA encod
13	1182.6	59.1	11058	6	AAD37229	Aad37229 Human dys
14	1182.6	59.1	11241	6	ABK82005	Abk82005 cDNA enco
15	1182.6	59.1	11443	6	ABK82002	Abk82002 DNA encod
16	1182.6	59.1	12923	1	AAN90338	Aan90338 Sequence
17	1182.6	59.1	13957	6	ABK81959	Abk81959 cDNA enco
18	1182.6	59.1	13957	6	ABT10904	Abt10904 Human bre
19	1182.6	59.1	13957	6	ABN95786	Abn95786 Gene #228
20	1182.6	59.1	13957	6	ABS69900	Abs69900 Human dys
21	1182.6	59.1	13977	6	ABS70403	Abs70403 Human bon
22	1182	59.1	2169	6	AAD37232	Aad37232 Human dys
23	1182	59.1	3531	6	AAD37238	Aad37238 Human and Becker muscular
24	1182	59.1	4498	6	AAD37258	Aad37258 Adeno-ass
25	1180	59.0	5339	6	ABK81998	Abk81998 DNA encod
26	1013	50.6	3510	6	AAD37240	Aad37240 Human dys
27	1013	50.6	4476	6	AAD37259	Aad37259 Adeno-ass
28	1004	50.2	13815	2	AAV18885	Aav18885 Mus muscu
29	1004	50.2	13815	6	ABK81960	Abk81960 cDNA enco
30	1004	50.2	13815	6	ABI99799	Abi99799 Mouse isc
31	1004	50.2	19307	6	AAT27558	Aat27558 Shuttle v
32	835	41.7	1821	6	AAD37241	Aad37241 Human dys
33	727	36.3	5417	6	ABK81997	Abk81997 DNA encod
34	678.4	33.9	4402	3	AAZ48568	Aaz48568 A rod sho
35	678.4	33.9	4414	6	AAD37260	Aad37260 Adeno-ass
36	678	33.9	1991	6	AAD37231	Aad37231 Human dys
37	677	33.8	1667	6	AAD37235	Aad37235 Human dys
38	666.4	33.3	3446	6	AAD37242	Aad37242 Human dys
39	618.8	30.9	4075	3	AAZ48569	Aaz48569 A rod sho
40	614.2	30.7	4402	3	AAZ48567	Aaz48567 A rod sho
41	537.8	26.9	3275	1	AAN97129	Aan97129 Partial s
42	448	22.4	1434	6	AAD37243	Aad37243 Human dys
43	408.2	20.4	11096	6	ABK81962	Abk81962 cDNA enco
44	408	20.4	10705	7	ABT41896	Abt41896 Toxicity
45	397	19.8	6045	2	AAT74665	Aat74665 Utrophin

ALIGNMENTS

```
RESULT 1
AAD37234
ID   AAD37234 standard; DNA; 3999 BP.
XX
AC   AAD37234;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Human dystrophin minigene delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 46-47; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
```

```
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
XX   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
SQ   Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;

Query Match             100.0%;  Score 2001;  DB 6;  Length 3999;
Best Local Similarity   100.0%;  Pred. NO. 0;
Matches 2001;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

Qy   1 GGCAGTTCATTGATGGGGAGTGAAGTAAACCTGGACCGTATCAAACAGTTTAGAAGAA   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1000 GGCAGTTCATTGATGGGGAGTGAAGTAAACCTGGACCGTATCAAACAGTTTAGAAGAA 1059

Qy   61 GTATTATCGTGGCTCTTTCTGCTGAGACACATTGCAAGCACAACAAGGAGAGAATTTCTAAT   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1060 GTATTATCGTGGCTCTTTCTGCTGAGACACATTGCAAGCACAACAAGGAGAGAATTTCTAAT 1119

Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACA   180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1120 GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACA 1179

Qy   181 GCCCATCAGGCGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1180 GCCCATCAGGCGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1239

Qy   241 AAATTATCAGAGATGAAGTACAGAAGACAGATGAATCTCCTAAATTCAAGA   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1240 AAATTATCAGAGATGAAGTACAGAAGACAGATGAATCTCCTAAATTCAAGA 1299

Qy   301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATG   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1300 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATG 1359

Qy   361 GATCTCCAGAATCAGAAACTGAAAGAGAGCCCTCAAGATGACTGGAACCTAAAACGCCAGTA   420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1360 GATCTCCAGAATCAGAAACTGAAAGAGAGCCCTCAAGATGACTGGAACCTAAAACGCCAGTA 1419

Qy   421 ACAAGGAAAATGAGGAGGAGCCTCTCAAGATGTCTTGAAGACCTCAGGGGTCAGGGTCAAGTA   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1420 ACAAGGAAAATGAGGAGGAGCCTCTCAAGATGTCTTGAAGACCTCAGGGGTCAGGGTCAAGTA 1479

Qy   481 CAACAACTAAGGTGCTTCAAGAGAATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1480 CAACAACTAAGGTGCTTCAAGAGAATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 1539

Qy   541 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTTGGAA   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGTTGGAA 1599

Qy   601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGC   660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1600 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGC 1659

Qy   661 TGGGTTCTTTTACAAGACCAGCCAGCTGACCTGACACTATTGGAGCCTCT   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1660 TGGGTTCTTTTACAAGACCAGCCAGCTGACCTGACACTATTGGAGCCTCT 1719

Qy   721 CCTACTCGACTGTTACTCTGGTGACAACCTGTGGTTACTAAGGAAACTGCCATCTCC   780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1720 CCTACTCGACTGTTACTCTGGTGACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 1779

Qy   781 AAACTAGAAATGCCATCTCGATGTTGGAGGTACCTACCCATAGATTACTGCAACAG   840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1780 AAACTAGAAATGCCATCTCGATGTTGGAGGTACCTACCCATAGATTACTGCAACAG 1839

Qy   841 TTCCCCCTTGGACACCTGGAAAAGTTTCTTGCCTGGATTCTTACAGGAGAGGAGTCTGAAAACAACTGCCAAT   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1840 TTCCCCCTTGGACACCTGGAAAAGTTTCTTGCCTGGATTCTTACAGGAGAGGAGTCTGAAAACAACTGCCAAT 1899

Qy   901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCAAGGGAGTAAAAGAG   960
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1900 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCAAGGGAGTAAAAGAG 1959

Qy   961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1960 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 2019

Qy   1021 AACCTGATGAAAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCGATGATGCAGTC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2020 AACCTGATGAAAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCGATGATGCAGTC 2079

Qy   1081 CTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2080 CTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 2139

Qy   1141 CTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGCTGGAAGCGTCTGCACCTTTCT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2140 CTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGCTGGAAGCGTCTGCACCTTTCT 2199

Qy   1201 CTGCAGGAACTTCTGTGTGGTGGCTACAGCTGAAAGATGATGATGAATTAAGCCGGCAGGCCACCT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2200 CTGCAGGAACTTCTGTGTGGTGGCTACAGCTGAAAGATGATGATGAATTAAGCCGGCAGGCCACCT 2259

Qy   1261 ATTGGAGGCGACTTTCCAGCAGTTCAGAGCAGCAGAACGATGTACATAGGGGCCTTCAAGAGG 1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2260 ATTGGAGGCGACTTTCCAGCAGTTCAGAGCAGCAGAACGATGTACATAGGGGCCTTCAAGAGG 2319

Qy   1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTAATCATGAGAAACTCTACCAGGAGACTACTCCTG 1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2320 GAATTGAAAACTAAAGAACCTGTAATCATGAGTAATCATGAGAAACTCTACCAGGAGACTACTCCTG 2379

Qy   1381 ACAGAGCAGCTTGGAGGACTAGAGAAACTCTACCAGGAGACTACTCCTCCT 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2380 ACAGAGCAGCTTGGAGGACTAGAGAAACTCTACCAGGAGACTACTCCTCCT 2439

Qy   1441 GAGGAGGAGCCCAGAATGTCACTCGGCTTCACGAAAGCAGCTGAGGAGGTCAATACT 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2440 GAGGAGGAGCCCAGAATGTCACTCGGCTTCACGAAAGCAGCTGAGGAGGTCAATACT 2499

Qy   1501 GAGTGTGGGAAAAATTGAACCTGCACTCCGCGACTCGGCAGCAGAAAAATAGAGTGAGACCCCTT 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2500 GAGTGTGGGAAAAATTGAACCTGCACTCCGCGACTCGGCAGCAGAAAAATAGAGTGAGACCCCTT 2559

Qy   1561 GAAAGACTTCCAGGAACTTCAAGAGGGCCACGATGAGCTGCGCGCAACTTCCTCCATTGAGCC 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2560 GAAAGACTTCCAGGAACTTCAAGAGGGCCACGATGAGCTGCGCGCAACTTCCTCCATTGAGCC 2619

Qy   1621 GAGGTGATCAAGGGATCCTGGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGAT 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2620 GAGGTGATCAAGGGATCCTGGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGAT 2679

Qy   1681 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGC 1740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2680 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGC 2739

Qy   1741 CAGGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGCATTCGGGGCATTGGGCATTCAGCTCTCCACTATAAC 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2740 CAGGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGCATTCGGGGCATTGGGCATTCAGCTCTCCACTATAAC 2799

Qy   1801 CTCAGCACTCTCGAAGAACCTGAACACCAGAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 1860
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2800 CTCAGCACTCTCGAAGAACCTGAACACCAGAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 2859

Qy   1861 CGAGTCAGGCAGCTGCATGAAGCTCGCGAGCTTACCACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2860 CGAGTCAGGCAGCTGCATGAAGCTCGCGAGCTTACCACTTTGGTCCAGCATCTCAGCACTTTCTT 2919

Qy   1921 TCCACGTCTGTCCAGGGTCCCTGGGAAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2920 TCCACGTCTGTCCAGGGTCCCTGGGAAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 2979

Qy   1981 ATCAACCACGAGACTCAAACA 2001
            |||||||||||||||||||||
Db   2980 ATCAACCACGAGACTCAAACA 3000

RESULT 2
AD37256
AD37256 standard; DNA; 4966 BP.
AAD37256;
21-AUG-2002 (first entry)
Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.
Homo sapiens.
Unidentified.
Chimeric.
WO200183695-A2.
08-NOV-2001.
27-APR-2001; 2001WO-US013677.
28-APR-2000; 2000US-0200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
gene.
Example 1; Page 59-60; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified N-
terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is AAV
vector plasmid construct containing human dystrophin minigenes, a muscle
creatine kinase (MCK) promoter and a small polyA signal sequence
Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;

Query Match   100.0%; Score 2001; DB 6; Length 4966;
Best Local Similarity 100.0%;   Pred. No.0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db   1937  GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA  1996
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241  AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCCTCTAAATTCAAGA   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1997  AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCCTCTAAATTCAAGA  2056

Qy    301  TGGGAATGCCTCAGGTAGCTAGCATGGGAAAACAAAAGCAATTACATAGAGTTTTAATG   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2057  TGGGAATGCCTCAGGTAGCTAGCATGGGAAAACAAAAGCAATTACATAGAGTTTTAATG  2116

Qy    361  GATCTCCAGAATCAGAAATCTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2117  GATCTCCAGAATCAGAAATCTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA  2176

Qy    421  ACAAGGAAAATGGAGGAAGAGCCTTGGACCTGATCTTGAACAACAAGACCTAAAACGCCAAGTA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2177  ACAAGGAAAATGGAGGAAGAGCCTTGGACCTGATCTTGAACAACAAGACCTAAAACGCCAAGTA  2236

Qy    481  CAACAACATAAGGTGCTTCAGGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2237  CAACAACATAAGGTGCTTCAGGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC  2296

Qy    541  ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAATCTAGTGGAGATCACCGCAACTGCTGCTTTGAA   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2297  ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGTGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTGAA  2356

Qy    601  GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2357  GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGC  2416

Qy    661  TGGGTTCTTTTACAAGACCAGCCTAGCTCCTGGACTGACCTATTGGAGCCTCT   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2417  TGGGTTCTTTTACAAGACCAGCCTAGCTCCTGGACTGACCTACCACCATTGGAGCCTCT  2476

Qy    721  CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAACTGCCATCTCC   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2477  CCTACTCAGACTGTTACTCTGTGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC  2536

Qy    781  AAACTAGAAATGCCATCTCCATTGATGTTGGAGTACCTACTCCATAGATTACTGCAACAG   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2537  AAACTAGAAATGCCATCTCCATTGATGTTGGAGTACCTACTCCATAGATTACTGCAACAG  2596

Qy    841  TTCCCCCTGACCTGGAAAAGTTTCTTCCTGGCTTACAGAGTCGAAACAACTGCCAAT   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2597  TTCCCCCTGACCTGGAAAAGTTTCTTCCTGGCTTACAGAAGTCGAAACAACTGCCAAT  2656

Qy    901  GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAG   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2657  GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAG  2716

Qy    961  CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGAGTTTATCAC  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2717  CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGAGTTTATCAC  2776

Qy   1021  AACCTGGATGAAACAGCCAAAAAATCTGAGATCCCTGAAGGTTCCGATGATGCAGTC  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2777  AACCTGGATGAAACAGCCAAAAAATCCTGAAGATCCCTGAAGGTTCCGATGATGCAGTC  2836

Qy   1081  CTGTTACAAAGACGTTGGATAACATCAACTTCAAGTGGAGTGAACTTCGGGAAAAAGTCT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2837  CTGTTACAAAGACGTTGGATAACATCAACTTCAAGTGGAGTGAACTTCGGGAAAAAGTCT  2896

Qy   1141  CTCAACATTAGGTGCCATTTGGAAGCCAGTTCTGAGAAGCGCTCGCACCCTTTCT  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2897  CTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGCTCGCACCCTTTCT  2956

Qy   1201  CTGCAGGAACTTCTGGTGTGGTACTACAGTTGAAAGATGATGAATTAAGCCGGCAGGACCT  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2957  CTGCAGGAACTTCTGGTGTGGTGCTACAGTTGAAAGATGATGAATTAAGCCGGCAGGACCT  3016

Qy   1261  ATTGGAGGGACTTCCAGCAGTTCAGCAGTTGATATTCTACAATTGGGAAGTGATTGGAACAGGA  1320
```

Query Match

```
Y      1  GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1757  GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA  1816

y     61  GTATTATCGTTGCTGGCTCTTTCTGCTGAGGACACAGTTCAAGCACACAGGAGATTTCAAT   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||
b   1817  GTATTATCGTTGCTGGCTCTTTCTGCTGAGGACACAGTTCAAGCACACAGGAGATTTCAAT  1876

y    121  GATGTGGAAGTGGTGAAAGACCAGTTCATATCTCATGAACATGATGATTTGACA   180
          |||||||||||||||||||||||||||||||||||||||||||
b   1877  GATGTGGAAGTGGTGAAAGACCAGTTCATATCTCATGAGGGGTACATGATGATTTGACA  1936

y    181  GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA   240
```

XX
PP　27-APR-2001; 2001WO-US013677.
XX
PR　28-APR-2000; 2000US-0200777P.
XX
PA　(XIAO/) XIAO X.
XX
PI　Xiao X;
XX
DR　WPI; 2002-049342/06.
XX
PT　New dystrophin minigene for treating Duchenne or Becker muscular
PT　dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT　rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT　gene.
XX
PS　Example 1; Page 67-68; 71pp; English.
XX
CC　The present invention relates to an isolated nucleotide sequence encoding
CC　a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC　terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC　domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC　invention also relates to a recombinant adeno-associated virus (AAV)
CC　comprising dystrophin minigene operably linked to an expression control
CC　element. The dystrophin minigene in operable linkage with an expression
CC　control element, in a recombinant adeno-associated virus or retrovirus is
CC　useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC　dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC　vector plasmid construct containing human dystrophin minigenes, a
CC　cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
SQ　Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;

Query Match　100.0%;　Score 2001;　DB 6;　Length 4990;
Best Local Similarity　100.0%;
Matches 2001;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　　1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGAGTAAACCGTTATCAAACAGCTTTAGAAGAA 60
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1781 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGAGTAAACCGTTATCAAACAGCTTTAGAAGAA 1840

Qy　　　61 GTTATTATCGTGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1841 GTTATTATCGTGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1900

Qy　　121 GATGTGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGATTTGACA 180
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1901 GATGTGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGATTTGACA 1960

Qy　　181 GCCCATCAGGGCGCGGGTGTTGGTAATATTCTACAATTGGAGTAAGCTGATTGGAACAGGA 240
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1961 GCCCATCAGGGCGCGGGTGTTGGTAATATTCTACAATTGGAGTAAGCTGATTGGAACAGGA 2020

Qy　　241 AAATTATCAGAAGATGAAGAAACTGAAGTACAGAGCCAGATGAGATCTCCTAAATTCAAGA 300
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2021 AAATTATCAGAAGATGAAGAAACTGAAGTACAGAGCCAGATGAGATCTCCTAAATTCAAGA 2080

Qy　　301 TGGGAATGCCTCAGGGTAGCCAGGGTAGCATGATGCTGAAAAACAAAGCAATTTACATAGAGTTTAATG 360
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2081 TGGGAATGCCTCAGGGTAGCCAGGGTAGCATGATGCTGAAAAACAAAGCAATTTACATAGAGTTTAATG 2140

Qy　　361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAAGA 420
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2141 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAAGA 2200

Qy　　421 ACAAGGAAAATGGAGGAAGAGCCCTTCTTGGACCTAAAACCGCCAAGTA 480
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||
Db　2201 ACAAGGAAAATGGAGGAAGAGCCCTTCTTGGACCTAAAACCGCCAAGTA 2260

Qy　　481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCTCTC 540
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2261 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCTCTC 2320

── (left column) ──

Db　3017 ATTGGAGGCGACTTTCCAGCAGTCAGTTCAGAAGCAGCAGAACGAGATGGTACATAGGGCCTTCAAGAGG 3076

Qy　1321 GAATTGAAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3077 GAATTGAAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3136

Qy　1381 ACAGAGAGCCCTTTGGAAGGACTAGAGAAACTTCTACCAGGACCCCAGAGAGCTGCCTCCT 1440
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3137 ACAGAGAGCAGCCCTTTGGAAGGACTAGAGAAACTTCTACCAGGACCCCAGAGAGCTGCCTCCT 3196

Qy　1441 GAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAAGCAGGTGAGGAGGTCAATACT 1500
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3197 GAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACT 3256

Qy　1501 GAGTGGGAAAAAATTGAACCTGACACTCCGCTGACTGCGCAGAGAAAAAATAGATGAGACCCCTT 1560
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3257 GAGTGGGAAAAAATTGAACCTGACACTCCGCTGACTGCGCAGAGAAAAAATAGATGAGACCCCTT 3316

Qy　1561 GAAAGACTCCAGGAACTTCAAGAGCTGGACGATGAGCTGGACCTGCGCCCAAGCT 1620
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3317 GAAAGACTCCAGGAACTTCAAGAGCTGGACGATGAGCTGGACCTGCGCCCAAGCT 3376

Qy　1621 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCTTCCTCAGGATCTCTCCCAAGAT 1680
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3377 GAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCTTCCTCAGGATCTCTCCCAAGAT 3436

Qy　1681 CACCTCGAGAAAGTCAAGGCACTTCAGGAGAAAATTGCGCCTCGAAAGAGAACGTGAGC 1740
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3437 CACCTCGAGAAAGTCAAGGCACTTCAGGAGAAAATTGCGCCTCGAAAGAGAACGTGAGC 3496

Qy　1741 CACGTCAATGACCTTGCTCGCGCCAGTTACCACTTTGGGCATTCAGCTCTCCACCGTATAAC 1800
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3497 CACGTCAATGACCTTGCTCGCGCCAGTTACCACTTTGGGCATTCAGCTCTCCACCGTATAAC 3556

Qy　1801 CTTCAGCACTCTGGAAGCACCTGCTGGAAGACTTCGCGAGGTTCTCGCAGGTGGCCGTCGAGGAC 1860
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3557 CTTCAGCACTCTGGAAGCACCTGGAAGACTTCGCGAGGTTCTCGCAGGTGGCCGTGCTGGAGGAC 3616

Qy　1861 CGAGTCAGGCAGCTGCAGTGCATGAAGCCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3617 CGAGTCAGGCAGCTGCAGTGCATGAAGCCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTT 3676

Qy　1921 TCCACGTCTGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　3677 TCCACGTCTGTCCAGGTCCCTGGGAGAGCAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3736

Qy　1981 ATCAACCACGAGACTCAAACA 2001
　　　　　|||||||||||||||||||||
Db　3737 ATCAACCACGAGACTCAAACA 3757

RSULT 3
AD37262
D　AAD37262 standard; DNA; 4990 BP.
XX
C　AAD37262;
XX
T　21-AUG-2002 (first entry)
XX
DE　Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
XX
GW　Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
GW　adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
GW　Becker muscular dystrophy; ds.
XX
OS　Homo sapiens.
OS　Cytomegalovirus.
OS　Unidentified.
OS　Chimeric.
XX
PN　WO20018369S-A2.
XX
PD　08-NOV-2001.

```
 541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA  600
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2321 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2380

 601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC  660
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2381 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 2440

 661 TGGGTTCTTTTACAAGACCAGCCTGACCTGACTGACCACTATTGGAGCCTCT          720
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
2441 TGGGTTCTTTTACAAGACCAGCCTGACCTGACTGACCACTATTGGAGCCTCT 2500

 721 CCTACTCAGACTGTTACTCTGGTGACAACAACCTGTCGTTACTAAGGAAAACTGCCATCTCC  780
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2501 CCTACTCAGACTGTTACTCTGGTGACAACAACCTGTCGTTACTAAGGAAAACTGCCATCTCC 2560

 781 AAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG  840
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2561 AAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 2620

 841 TTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT  900
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2621 TTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 2680

 901 GTCCTACAGGATGCTACCCGTAAGGAAGAGGCTCCTAGAGACTCCAAGGGAGTAAAAGAG  960
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2681 GTCCTACAGGATGCTACCCGTAAGGAAGAGGCTCCTAGAGACTCCAAGGGAGTAAAAGAG 2740

 961 CTGATGAAACAATGGCAAGAACTCCAAGGTGAAATTGAAGCTCACCACGATGTTTATCAC 1020
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2741 CTGATGAAACAATGGCAAGAACTCCAAGGTGAAATTGAAGCTCACCACGATGTTTATCAC 2800

1021 AACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGTTCCTGAAGTTCCTGATGATGCAGTC 1080
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2801 AACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGTTCCTGAAGTTCCTGATGATGCAGTC 2860

1081 CTGTTACAAAGACGTTTGGATAACAATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 1140
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2861 CTGTTACAAAGACGTTTGGATAACAATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 2920

1141 CTCAACATTAGGTCCCATTGGAAGCCCAGTTCTGACCAGTGGAAGCGTCTTGCACCTTCT 1200
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2921 CTCAACATTAGGTCCCATTGGAAGCCCAGTTCTGACCAGTGGAAGCGTCTTGCACCTTCT 2980

1201 CTGCAGGAACTTCTGGTGGTGGCTACACGTTAAGCCCGGCCAGGCACCT 1260
     ||||||||||||||||||||||||||||||||||||||||||||||||
2981 CTGCAGGAACTTCTGGTGGTGGCTACACGTTAAGCCCGGCCAGGCACCT 3040

1261 ATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3041 ATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 3100

1321 GAATTGAAACTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGTACGATATTTCTG 1380
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3101 GAATTGAAACTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGTACGATATTTCTG 3160

1381 ACAGAGCAGCCCTTTGGAAGGACTAGAGAAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCT 1440
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3161 ACAGAGCAGCCCTTTGGAAGGACTAGAGAAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCT 3220

1441 GAGGAGAGCCCAGAGCCACCAGGATGAGCTGGAGGTCAATACT 1500
     |||||||||||||||||||||||||||||||||||||||||||
3221 GAGGAGAGCCCAGAGCCACCAGGATGAGCTGGAGGTCAATACT 3280

1501 GAGTGGGAAAAATTGAACCTGCACTCCGCTACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTT 1560
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3281 GAGTGGGAAAAATTGAACCTGCACTCCGCTACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTT 3340

1561 GAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTGCGCCGCCAAGCT 1620
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3341 GAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTGCGCCGCCAAGCT 3400

1621 GAGGTGATCAAGGGGATCCTGGCAGCCCGTGGGCGATCTCCATTGACTCTCTCCAAGAT 1680
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3401 GAGGTGATCAAGGATCCTGGCAGCCCGTGGGCGATCTCCATTGACTCTCTCCAAGAT 3460

1681 CACCTCGAGAAGTCAAGGCACTTCGAGAGGAAATTGCGCCTCTGAAAGAGAACGTGAGC 1740
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3461 CACCTCGAGAAGTCAAGGCACTTCGAGAGGAAATTGCGCCTCTGAAAGAGAACGTGAGC 3520

1741 CACGTCAATGACCTTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 1800
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3521 CACGTCAATGACCTTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 3580

1801 CTCAGCACTCCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 1860
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3581 CTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 3640

1861 CGAGTCAGGCAGCGACTGCATGAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTT 1920
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3641 CGAGTCAGGCAGCGACTGCATGAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTT 3700

1921 TCCACGTCTGTCCAGGGTCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3701 TCCACGTCTGTCCAGGGTCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3760

1981 ATCAACCACGAGACTCAAACA 2001
     |||||||||||||||||||||
3761 ATCAACCACGAGACTCAAACA 3781
```

RESULT 4
AAD37237
ID   AAD37237 standard; DNA; 3858 BP.
XX
AC   AAD37237;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 48-49; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is

```
:C  useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
:C  dystrophy (BMD) in a mammalian subject. The present sequence is human
:C  dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
:C  hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
:X  and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
:Q  Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;

Query Match  85.4%;  Score 1709;  DB 6;  Length 3858;
Best Local Similarity  93.0%;  Pred. No. 0;
Matches 1860;  Conservative  0;  Mismatches  0;  Indels  141;  Gaps  1;
```

```
Qy    1  GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAA    60
Db 1000  GGCAGTTCATTGATGATGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAA 1059

Qy   61  GTATTATCGTGCTCTTCTTCTGCTTGCTGAGGACACCAGTCAAGCACACAGGAGATTTCTAAT   120
Db 1060  GTATTATCGTGCTCTTCTTCTGCTTGCTGAGGACACCAGTCAAGCACACAGGAGAGATTTCTAAT 1119

Qy  121  GATGTGAAGTGGTGAAAGACCAGTTTCATCTCATGAGGGTACATGATGAGATTTGACA   180
Db 1120  GATGTGAAGTGGTGAAAGACCAGTTTCATCTCATGAGGGTACATGATGAGATTTGACA 1179

Qy  181  GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGA   240
Db 1180  GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGA 1239

Qy  241  AAATTATCAGAGAAGATGAAGAAAACTGAAGTACAAGACAGATGAATCTGAATCTCCTAAATTCAGA   300
Db 1240  AAATTATCAGAGAAGATGAAGAAAACTGAAGTACAAGACAGATGAATCTGAATCTCCTAAATTCAGA 1299

Qy  301  TGGGAATGCCTAGCAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATG   360
Db 1300  TGGGAATGCCTAGCAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 1359

Qy  361  GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA   420
Db 1360  GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 1419

Qy  421  ACAAGGAAAATGGAGGAGGAGAGCCTCTTGACCTGATCTTTGAAGACCTAAAACGGCCAAGTA   480
Db 1420  ACAAGGAAAATGGAGGAGGAGAGCCTCTTGACCTGATCTTTGAAGACCTAAAACGGCCAAGTA 1479

Qy  481  CAACAACATAAGGTGCTTCAAGAGATCTAGAAGACAAGAACAAGTCAAGCTGCTTGCTGGAA   540
Db 1480  CAACAACATAAGGTGCTTCAAGAGATCTAGAAGACAAGAACAAGTCAAGCTGCTTGCTGGAA 1539

Qy  541  ACTCACATGGTGGTAGTGGTAGTTCAGTGAATCAGTGGAGATCAGTGGCAAACATCTGAA   600
Db 1540  ACTCACATGGTGGTAGTGGTAGTTCAGTGAATCAGTGGAGATCAGTGGCAAACATCTGAA 1599

Qy  601  GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC   660
Db 1600  GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 1659

Qy  661  TGGGTTCTTTACAAGACCAGCTGAAGACCTGACCACTATTGGAGCCTCT   720
Db 1660  TGGGTTCTTTACAAGAC---------------------------- 1677

Qy  721  CCTACTCAGACTGTTACTCTGTGACACCAGTGGTGGTTACTAAGGAAACTGCCATCTCC   780
Db 1678  -------------------------------------- 1677

Qy  781  AAACTAGAAAATGCCATCTTCCTTGATGTTGGAGGTACCACTCATGATTACTGCAACAG   840
Db 1678  ----------ACTCATGATTACTTGCAACAG 1698

Qy  841  TTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGGCTGAAACAACTGCCAAT   900
Db 1699  TTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGGCTGAAACAACTGCCAAT 1758

Qy  901  GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGAACTCCAAGGGAGTAAAGAG   960
```

```
Db 1759  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGAACTCCAAGGGAGTAAAAGAG 1818

Qy  961  CTGATGAAACAATGGCAAGAGACCTCCAAGGTGAAATTGAAGTTGAAGTTGAGCTCACACAGATGTTTATCAC 1020
Db 1819  CTGATGAAACAATGGCAAGAGACCTCCAAGGTGAAATTGAAGTTGAAGTTGAGCTCACACAGATGTTTATCAC 1878

Qy 1021  AACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCTCCTGAGAGTTCCGATGATGCAGTC 1080
Db 1879  AACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCTCCTGAGAGTTCCGATGATGCAGTC 1938

Qy 1081  CTGTTACAAAGACGTTGGATAACATGAACCTCAAGTGGATGTGGAACTTCGGAAAAAGTCT 1140
Db 1939  CTGTTACAAAGACGTTGGATAACATGAACCTCAAGTGGATGTGGAACTTCGGAAAAAGTCT 1998

Qy 1141  CTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCT 1200
Db 1999  CTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCT 2058

Qy 1201  CTGCAGGAACTTCTGTGTCGTGGCTACAGCTGGAGGGTACAGCTGTACATGATGAGATTGACA 1260
Db 2059  CTGCAGGAACTTCTGTGTCGTGGCTACAGCTGGAGGGTACAGCTGTACATGATGAGATTGACA 2118

Qy 1261  ATTGGAGGCGACTTTCCAGAGTTCAGAGTCAGAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320
Db 2119  ATTGGAGGCGACTTTCCAGAGTTCAGAGTCAGAGCAGAACGATGTACATAGGGCCTTCAAGAGG 2178

Qy 1321  GAATTGAAAACTAAAGAACTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
Db 2179  GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 2238

Qy 1381  ACAGAGCAGCCTTTGAAGGACTAGAGAACTCTACCAGGAGCCCAGAGAGCTGCCTCCT 1440
Db 2239  ACAGAGCAGCCTTTGAAGGACTAGAGAACTCTACCAGGAGCCCAGAGAGCTGCCTCCT 2298

Qy 1441  GAGGAGAGAGCCCAGAATGTCACTCGGCTTCCGCTGACTGGCAGAGAAAGCAGGGTCAATACT 1500
Db 2299  GAGGAGAGAGCCCAGAATGTCACTCGGCTTCCGCTGACTGGCAGAGAAAGCAGGGTCAATACT 2358

Qy 1501  GAGTGGGAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTT 1560
Db 2359  GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTT 2418

Qy 1561  GAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGCGCCAAGCTTCCTCCCAAGAT 1620
Db 2419  GAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGCGCCAAGCTTCCTCCCAAGAT 2478

Qy 1621  GAGGTGATCAAGGATCCTGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCTCCAGAT 1680
Db 2479  GAGGTGATCAAGGATCCTGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCTCCAAGAT 2538

Qy 1681  CACCTCGAGAAGTCAAGGCACTTCGAGCACTTCGAGGAAATTGGCCTCTGAAAGAGAACGTGAGC 1740
Db 2539  CACCTCGAGAAGTCAAGGCACTTCGAGCACTTCGAGGAAATTGGCCTCTGAAAGAGAACGTGAGC 2598

Qy 1741  CACGTCAATGACCTTCGCCGCCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 1800
Db 2599  CACGTCAATGACCTTCGCCGCCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 2658

Qy 1801  CTCAGCACTCTGGAAGACTCGATGAGAACACCAGTGAACACCAGGCTTGGTCCAGCACTTTCTT 1860
Db 2659  CTCAGCACTCTGGAAGACTCGATGAGAACACCAGTGAACACCAGGCTTGGTCCAGCACTTTCTT 2718

Qy 1861  CGAGTCAGGCAGCTGCATGAAGCCTCAGAGGACTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 2719  CGAGTCAGGCAGCTGCATGAAGCCTCAGAGGACTTGGTCCAGCATCTCAGCACTTTCTT 2778

Qy 1921  TCCACGTCTGTCCAGGTCCCTGGAGAGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
Db 2779  TCCACGTCTGTCCAGGTCCCTGGAGAGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 2838

Qy 1981  ATCAACCACGAGACTCAAACA 2001
```

b 2839 ATCAACCACCAGAGACTCAAACA 2859

ESULT 5
AD37257
D AAD37257 standard; DNA; 4825 BP.
X AAD37257;
X 21-AUG-2002 (first entry)
E Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
W Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.
X Homo sapiens.
S Unidentified.
S Chimeric.
X WO20183695-A2.
D 08-NOV-2001.
X 27-APR-2001; 2001WO-US013677.
F 28-APR-2000; 2000US-0200777P.
A (XIAO/) XIAO X.
A Xiao X;
I
R WPI; 2002-049342/06.
X
T New dystrophin minigene for treating Duchenne or Becker muscular
T dystrophy comprises an N-terminal domain or modified N-terminal domain,
T rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
T gene.
X
S Example 1; Page 61-62; 71pp; English.
X
C The present invention relates to an isolated nucleotide sequence encoding
C a dystrophin minigene. The minigene comprises N-terminal or modified N-
C terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
C domains and cysteine-rich domains of dystrophin or utrophin genes. The
C invention also relates to a recombinant adeno-associated virus (AAV)
C comprising dystrophin minigene operably linked to an expression control
C element. The dystrophin minigene in operable linkage with an expression
C control element, in a recombinant adeno-associated virus or retrovirus is
C useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
C dystrophy (BMD) in a mammalian subject. The present sequence is AAV
C vector plasmid construct containing human dystrophin minigenes, a muscle
C creatine kinase (MCK) promoter and a small polyA signal sequence.
X
Q Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;

Query Match 85.4%; Score 1709; DB 6; Length 4825;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

Qy 181 GCCCATCAGGCGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1937 GCCCATCAGGCGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGA 1996

Qy 241 AAATTATCAGAGATGAAGAAACTGAAGTACAGAGACAGATGAATCTCCTAAATTCAAGA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1997 AAATTATCAGAGATGAAGAAACTGAAGTACAGAGACAGATGAATCTCCTAAATTCAAGA 2056

Qy 301 TGGGAATGCTCAGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATG 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2057 TGGGAATGCCTCAGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATG 2116

Qy 361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2117 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 2176

Qy 421 ACAAGGAAAATGGAGGAAGAGCCCTTGGACCTGATCTTGAAGACCTAAAACCGCCAAGTA 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2177 ACAAGGAAAATGGAGGAAGAGCCCTTGGACCTGATCTTGAAGACCTAAAACCGCCAAGTA 2236

Qy 481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2237 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 2296

Qy 541 ACTCCACATGGTGGTGGTAGTTGATGAATCTAGTTGGATGAATCTAGTTGGATGCTTTGGAA 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2297 ACTCCACATGGTGGTGGTAGTTGATGAATCTAGTTGGATGAATCTAGTTGGATGCTTTGGAA 2356

Qy 601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGC 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2357 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGC 2416

Qy 661 TGGGTTCTTTTACAAGAC------------------------------------ 2434
        ||||||||||||||
Db 2417 TGGGTTCTTTTACAAGAC 2434

Qy 721 CCTACTCAGACTGTTACTCGGTGACACAACCTGGTTACTAAGGAAACTGCCATCTCC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2435 -- 2434

Qy 781 AAACTAGAAATGCCATCTTCCTTGATGTTGGAGTACCTACTCATAGATTACTGCAACAG 840
        ||||||||||||||||||||||
Db 2435 ---------------ACTCATAGATTACTGCAACAG 2455

Qy 841 TTCCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2456 TTCCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 2515

Qy 901 GTCCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCAAGGGAGTAAAAGAG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2516 GTCCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCAAGGGAGTAAAAGAG 2575

Qy 961 CTGATGAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAC 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2576 CTGATGAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAC 2635

Qy 1021 AACCTGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2636 AACCTGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTC 2695

Qy 1081 CTGTTACAAAGACGTTTGATAACATGACTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2696 CTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 2755

Qy 1141 CTCAACATTAGGTCCCATTGGAGCCAGTTCTGACCAGTGGAGCGGTCTGCACCTTTCT 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2756 CTCAACATTAGGTCCCATTGGAGCCAGTTCTGACCAGTGGAGCGGTCTGCACCTTTCT 2815

Qy 1201 CTGCAGGAACTTCGTGTGTGGCTACAGCTGAAAGATGAATGAATGATGGACCT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2816 CTGCAGGAACTTCGTGTGTGGCTACAGCTGAAAGATGAATGAATGATGGCACCT 2875

Qy 1261 ATTGGAGGCGACTTTCAGCAGTTCAGAGACAGAACGATGTACATAGGGCCTTCAAGAGG 1320

Qy 1 GGCAGTTCATTGATGGATGGAGTGAAGTAAACCTGACCGTTATCAAAACAGCTTTAGGAGAA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b 1757 GGCAGTTCATTGATGGATGGAGTGAAGTAAACCTGACCGTTATCAAAACAGCTTTAGAGAA 1816

Qy 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCCAAGCACAAGGAGCACAGATTCTAAT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b 1817 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCCAAGCACAAGGAGCACAGATTCTAAT 1876

Qy 121 GATGTGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGATTTGACA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b 1877 GATGTGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGATTTGACA 1936

PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 68-70; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;

Query Match   85.4%; Score 1709; DB 6; Length 4848;
Best Local Similarity   93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

Qy   1 GGCAGTTCATTGATGGAGAGTGAAGTAAACTGGACGGCCGTTATCAAACAGCTTAGAAGA   60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1780 GGCAGTTCATTGATGGAGAGTGAAGTAAACTGGACGGCCGTTATCAAACAGCTTAGAAGA 1839

Qy   61 GTATTATCGTGGCTTCTTCGCTGGAGACACATTGCAAGCACACAAGGAGAGATTTCTAAT 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1840 GTATTATCGTGGCTTCTTCGCTGGAGACACATTGCAAGCACACAAGGAGAGATTTCTAAT 1899

Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACA 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1900 GATGTGGAAGTGGTGAAAGACCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1959

Qy   181 GCCCATCAGGGCGCGGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1960 GCCCATCAGGGCGCGGGTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 2019

Qy   241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGAGAGATGAATCTCCTAAATTCAAGA 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2020 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGAGAGATCTCCTAAATTCAAGA 2079

Qy   301 TGGGAATGCCTTCAGGGTAGCTAGCAGGGTCATGGAAAAACAAGCAATTTACATAGAGTTTTAATG 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2080 TGGGAATGCCTTCAGGGTAGCTAGCAGGGTCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATG 2139

Qy   361 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAGA 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2140 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAGA 2199

Qy   421 ACAAGGAAAATGGAGAAGAGCCCTTGACCTAAGACTAAAACGCCAAGTA 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2200 ACAAGAAAATGGAGAAGAGCCCTTGACCTAAGACTAAAACGCCAAGTA 2259

Qy   481 CAACAACATAAGGTGCTTCAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2260 CAACAACATAAGGTGCTTCAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 2319

Db   2875 ATTGGAGGGACTTCCAGCAGTTCAGAGAGCAGAACGATGTACATAGGGCCTTCAAGAGG
Qy   1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
Db   2936 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 2995

Qy   1381 ACAGAGCAGCCTTTGGAAGGACTAGAGAAAACTTCACCAGGAGACGTGCCTCCT 1440
Db   2996 ACAGAGCAGCCTTTGGAAGGACTAGAGAAAACTTCACCAGGAGACGTGCCTCCT 3055

Qy   1441 GAGGAGAGAGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGTCAATACT 1500
Db   3056 GAGGAGAGAGCCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGTCAATACT 3115

Qy   1501 GAGTGGGAAAAATTGAACCTGCACTCCGCTCCGTCACTCCGCTCCGCGAGAAAATAGATGAGACCCTT 1560
Db   3116 GAGTGGGAAAAATTGAACCTGCACTCCGCTCCGTCACTCCGCTCCGCGAGAAAATAGATGAGACCCTT 3175

Qy   1561 GAAAGACTCCAGAACTTCAAGAGGGCCACGGATGAGCTGCACTTCAAGCTCGCGCCAAGCT 1620
Db   3176 GAAAGACTCCAGAACTTCAAGAGGGCCACGGATGAGCTGCACTTCAAGCTCGCGCCAAGCT 3235

Qy   1621 GAGGTGATCAAGGGATCGCAGCCCGTGGGCGATCTCCTCATTGACTCTTCCCAAGAT 1680
Db   3236 GAGGTGATCAAGGGATCCTGGCGCCCGTGGGCGATCTCCTCATTGACTCTTCCCAAGAT 3295

Qy   1681 CACCTGAGAAAAGTCAAGCACTTCGAGGAGAAATTGCGCCCTCTGAAAAGAGAACGTGAGC 1740
Db   3296 CACCTGAGAAAAGTCAAGCACTTCGAGGAGAAATTGCGCCCTCTGAAAAGAGAACGTGAGC 3355

Qy   1741 CACGTCAATGACCTTGCTCGCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 1800
Db   3356 CACGTCAATGACCTTGCTCGCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 3415

Qy   1801 CTCAGCACTCTGGAAGAGCTGAACACCAGATGAAGCTTTCGTCCAGCCATCTCAGCACTTTCTT 1860
Db   3416 CTCAGCACTCTGGAAGAGCTGAACACCAGATGAAGCTTTCGTCCAGCCATCTCAGCACTTTCTT 3475

Qy   1861 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGTCCAGCCATCTCAGCACTTTCTT 1920
Db   3476 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGTCCAGCACTTTCCAGCACTTTCTT 3535

Qy   1921 TCCACGTCTGTCAGGGTCCCTGGAGAGAGCCATCTCGCCAAAACAAAGTCCCTACTAT 1980
Db   3536 TCCACGTCTGTCAGGGTCCCTGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTAT 3595

Qy   1981 ATCAACCACGAGACTCAAACA 2001
             |||||||||||||||||||||
Db   3596 ATCAACCACGAGACTCAAACA 3616

RESULT 6
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX
AC AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX

```
 541 ACTCACATGGTGGTGAGTTGATGAATCTAGTGAGATCACGGCAACTGCTGTTTGAA  600
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2320 ACTCACATGGTGGTGAGTTGATGAATCTAGTGAGATCACGGCAACTGCTGTTTGAA 2379

 601 GAACAACTTAAGGTATTGGGAGGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC  660
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2380 GAACAACTTAAGGTATTGGGAGGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 2439

 661 TGGGTTCTTTTACAAGACCAGCCAGCCTGAGCTCCTGGACTGACGACCACCATTGGAGCCTCT  720
     ||||||||||||||||
2440 TGGGTTCTTTTACAAGAC------                                       2457

 721 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTTACTAAGGAAACTGCCATCTCC  780
     ------------                                                   2457

 781 AAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG  840
     --------------ACTCATAGATTACTGCAACAG                           2478

 841 TTCCCCCTGGACCTGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT  900
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2479 TTCCCCCTGGACCTGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 2538

 901 GTCCTACAGGATGTGCTACCCGTAAGCAAGACTCCTCCAAGGAGACTTCCAAGGGAGTAAAAGAG  960
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2539 GTCCTACAGGATGTGCTACCCGTAAGCAAGACTCCTCCAAGGAGACTTCCAAGGGAGTAAAAGAG 2598

 961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACAGATGTTTATCAC 1020
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2599 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACAGATGTTTATCAC 2658

1021 AACCTGATGAAAACAGCCAAAAAATCCTGAGATCCTGAGATGATGCAGTC 1080
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2659 AACCTGATGAAAACAGCCAAAAAATCCTGAGATCCTGAGATGATGCAGTC 2718

1081 CTGTTACAAGACGTTTGGATAACATGAACTCAGTGGAGTGAACTTCGGAAAAAGTCT 1140
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2719 CTGTTACAAAGACGTTTGGATAACATGAACTCAGTGGAGTGAACTTCGGAAAAAGTCT 2778

1141 CTCAACACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCCACCTTTCT 1200
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2779 CTCAACACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCCACCTTTCT 2838

1201 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCCAGGCACCT 1260
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2839 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCCAGGCACCT 2898

1261 ATTGGAGGCGACTTTGGACACTTCAGAAGCCAGAACGAGGCCTTCAAGAGG 1320
     |||||||||||||||||||||||||||||||||||||||||||||||||||
2899 ATTGGAGGCGACTTTGGACACTTCAGAAGCCAGAACGAGGCCTTCAAGAGG 2958

1321 GAATTGAAAACTAAAGAACCTGTAATCATGATGACTGTACTCTTGAGACTGTACGAATATTCTG 1380
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2959 GAATTGAAAACTAAAGAACCTGTAATCATGATGACTGTACTCTTGAGACTGTACGAATATTCTG 3018

1381 ACAGAGGAGAGCCAGAGTCTTGGAAGACTAGAGAAACTCTACCAGGAGAACTGCCTCCT 1440
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3019 ACAGAGGAGAGCCAGAGTCTTGGAAGACTAGAGAAACTCTACCAGGAGAACTGCCTCCT 3078

1441 GAGGAGAGAGCCAGAGCTGCACTCACTGCCTTCTACGAAAAGCAGGTTCAATACT 1500
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
3079 GAGGAGAGAGCCAGAGCTGCACTCACTGCCTTCTACGAAAAGCAGGTTCAATACT 3138

1501 GAGTGGAAAAATTGAACCTGCACTCGCTGACTGCAGAGAAAAATAGATGAGACCCTT 1560
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3139 GAGTGGAAAAATTGAACCTGCACTCGCTGACTGCAGAGAAAAATAGATGAGACCCTT 3198

1561 GAAAGACTTCCAGGAACTTCAGAGAGCCACCGATGAGCTGGACTTCGCCAAGCT 1620
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
3199 GAAAGACTTCCAGGAACTTCAGAGAGCCACCGATGAGCTGGACTTCGCCAAGCT 3258
```

```
1621 GAGGTGATCAAGGATCCTGGCCAGCCGGTGGGCCATCCTCCTCATTGACTCTCCCAAGAT 1680
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3259 GAGGTGATCAAGGATCCTGGCCAGCCGGTGGGCCATCCTCCTCATTGACTCTCCCAAGAT 3318

1681 CACCTGAGAAAGTCAAGGACACTTCGGCCTCTGAGGAGAAATTGCGCGTTCTGAAAGAGAAGTGAGC 1740
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3319 CACCTGAGAAAGTCAAGGACACTTCGGCCTCTGAGGAGAAATTGCGCGTTCTGAAAGAGAAGTGAGC 3378

1741 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 1800
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3379 CACGTCAATGATGACCTCGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAT 3438

1801 CTCAGCACTCTGGAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGCGAGGAC 1860
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3439 CTCAGCACTCTGGAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGCGAGGAC 3498

1861 CGAGTCAGGCAGCTGCATGGAGCCCACAGGGACTTTGGTCCAGCATCCAGCACTTTCTT 1920
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3499 CGAGTCAGGCAGCTGCATGGAGCCCACAGGGACTTTGGTCCAGCATCCAGCACTTTCTT 3558

1921 TCCACGTCGTCCAGGTCCCTGGGAGAGAGCCATCTCCGCCAAACAAAGTGCCCTACTAT 1980
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3559 TCCACGTCGTCCAGGTCCCTGGGAGAGAGCCATCTCCGCCAAACAAAGTGCCCTACTAT 3618

1981 ATCAACCACGAGACTCAAACA 2001
     |||||||||||||||||||||
3619 ATCAACCACGAGACTCAAACA 3639
```

```
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a
XX   small polyA signal sequence
SQ   Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

Query Match       85.4%;  Score 1709;  DB 6;  Length 5060;
Best Local Similarity   93.0%;  Pred. No. 0;
Matches 1860;  Conservative   0;  Mismatches   0;  Indels 141;  Gaps   1;

Qy      1 GGCAGTTCATTGATGGAGAGTGAAGTAACCTGACCGTTATCAAACAGCTTTAGAAGAA      60
          |||||||||||||||||||||||||||     |||||||||||||||||||||||||||||
Db   1992 GGCAGTTCATTGATGGAGAGTGAAGTGAAGTAACCTGACCGTTATCAAACAGCTTTAGAAGAA   2051

Qy     61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT     120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2052 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT    2111

Qy    121 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGATGGGGTACATGATGATTTGACA     180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2112 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGATGGGGTACATGATGATTTGACA    2171

Qy    181 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGA     240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2172 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGA    2231

Qy    241 AAATTATCAGAAGATGAAGAATGAAGAAGACTAGAACAGAGACAGATGATGAATCCTAAATTCAAGA     300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2232 AAATTATCAGAAGATGAAGAAACTGAAGATGAAGAACAGATGAATCTCCTAAATTCAAGA    2291

Qy    301 TGGGAATGCCTCAGGGTAGCTAGCATGAAGGAAAAACAAAGCAATTTACATAGAGTTTTAATG     360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2292 TGGGAATGCCTCAGGGTAGCTAGCATGAAGGAAAAACAAAGCAATTTACATAGAGTTTTAATG    2351

Qy    361 GATCTTCCAGAATCAGAAAGAAGAGTTGAATGACTCGGCTAACAAACAGAAGAAAGA     420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2352 GATCTCCCAGAATCAGAAACTGAAAGAAGAGTTGAATGACTCGGCTAACAAACAGAAGAAAGA    2411

Qy    421 ACAAGGAAAAATGAGGAGGAAGAGCGCCTCTTGGAACCTGATCTTGAAGAACCGCCAAGTA     480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2412 ACAAGGAAAAATGAGGAGGAAGAGCGCCTCTTGGAACCTGATCTTGAAGAACGCGCCAAGTA    2471

Qy    481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC     540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2472 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC    2531

Qy    541 ACTCACACATGGTGGTGGTAGTTGATGAATCTAGTGGACAACACGCAACTGCTTTGGAA     600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2532 ACTCACACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTTTGGAA    2591

Qy    601 GAACAACTTAAGGTATTGGGAGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAAGAACCGC     660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2592 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAAGAACCGC    2651

Qy    661 TGGGTTCTTTTACAAGAGAC----------------------------------------     2669
          |||||||||||||||||||
Db   2652 TGGGTTCTTTTACAAGAC----------------------------------------    2669

Qy    721 CCTACTCCAGACTGTTACTCTGGTGACACAACCTGGTGTTACTAAGGAAACTGCCATCTCC     780
          ||||||||||||||||||||||||||||||||||
Db   2669 ----------------------------------------------------

Qy    781 AAACTAGAAAATGCCATCTTCCTTCGATGTTGGAGTGGGTACCTACTCATAGATTACTGCAACAG     840
          ||||||||||||||||||||||||||||||||||||||
Db   2670 --------------------------ACTCATAGATTACTGCAACAG    2690

Qy    841 TTCCCCCTGACCTGGAAAGTTTCTTGCCTGGCTGCCTTACAGAGACTGAAAACAACTGCCAAT     900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2691 TTCCCCCTGACCTGGAAAGTTTCTTGCCTGGCTGCCTTACAGAGACTGAAAACAACTGCCAAT    2750

Qy    901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTTCCTAAGAACTCCAAGGAGTAAAAGAG     960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2751 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTTCCTAAGAACTCCAAGGAGTAAAAGAG    2810

Qy    961 CTGATGAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAC     1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2811 CTGATGAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAC    2870

Qy   1021 AACCTGATGAAACAGCCAAAAAATCCTGAGATCCTGGATGATCCTGATGATGCAGTC     1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2871 AACCTGATGAAACAGCCAAAAAATCCTGAGATCCTGGATGATCCTGATGATGCAGTC    2930

Qy   1081 CTGTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAGTGGAGTGAACTTCGGAAAAGTCT     1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2931 CTGTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAGTGGAGTGAACTTCGGAAAAAGTCT    2990

Qy   1141 CTCAACATTAGGTCCCATTGGAAGCCCAGTTCTGACCAGTGGAGAGGCGTCTGCACTTTCT     1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2991 CTCAACATTAGGTCCCATTGGAAGCCCAGTTCTGACCAGTGGAGAGGCTGCTGCACCCTTTCT    3050

Qy   1201 CTGCAGGAACTCTCGGTGGTGTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCT     1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 CTGCAGGAACTCTCGGTGGTGTACAGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCT    3110

Qy   1261 ATTGGAGGGACTTTCCAGCAGTTCAGAGAGCAAGTAAGCTGATTGGAACGATGTACATAGAGG     1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 ATTGGAGGCGACTTTCCAGCAGTTCAGAGAGCAAGTAAGCTGATTGGAACGATGTACATAGAGG    3170

Qy   1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGTACGAGATATTTTCTG     1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3171 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTCTTGAGACTGTACGAGATATTTTCTG    3230

Qy   1381 ACAGCAGCCTTGGAGGAGCCCAGAATGTCACTCGGCTTCACTGAGAAACTCTACCAGGAGAGCTGCCTCCT     1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 ACAGCAGCCTTGGAGGAGCCCAGAATGTCACTCGGCTTCACTGAGAAACTCTACCAGGAGAGCTGCCTCCT    3290

Qy   1441 GAGGAGAGAGCCCAGAATGTCACTCGGCTTCACTGAGAAAGCAGGCTGAGGAGGTCAATACT     1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 GAGGAGAGAGCCCAGAATGTCACTCGGCTTCACTGAGAAAGCAGGCTGAGGAGGTCAATACT    3350

Qy   1501 GAGTGGGAAAAATTGAACCTGCTGACTGGCAGTGGCCAGAGAAAATAGATGAGACCCTT     1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3351 GAGTGGGAAAAAATTGAACCTGCTGACTGGCAGTGGCCAGAGAAAATAGATGAGACCCTT    3410

Qy   1561 GAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGCCCAAGCT     1620
          |||||||||||||||||||||||||||||||||||||||||||||
Db   3411 GAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGCGCTGCCCAAGCT    3470

Qy   1621 GAGGTGATCAAGGGATCTGGGCGGTGGCGGCGATCTCTCATTGACTTCTCTCCAAGAT     1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3471 GAGGTGATCAAGGGATCTGGGGCGGTGGCGGCGATCTCCTCATTGACTCCTCTCCAAGAT    3530

Qy   1681 CACCTCGGAAAGTCAAGGCACTTCAAGGCACTTCAGAGGCCTCTGAAAAGAGACGTGAGC     1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3531 CACCTCGGAAAGTCAAGGCACTTCAAGGCACTTCAGAGGCCTCTGAAAAGAGAACGTGAGC    3590

Qy   1741 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC     1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3591 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC    3650

Qy   1801 CTCAGCACTCTGGAAGACCTGCGATGGGAAGCTTCTGCAGGTGGCCGTGGCGTCGAGGAC     1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3651 CTCAGCACTCTGGAAGACCTGCGATGGGAAGCTTCTGCAGGTGGCCGTGGCGTCGAGGAC    3710

Qy   1861 CGAGTCAGGCAGCTGCATGAGGTGCATGAAGCGGGACTTTGGTCCAGCATCTCAGCACTTTCTT     1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3711 CGAGTCAGGCAGCTGCATGAGGTGCATGAAGCGGGACTTTGGTCCAGCATCTCAGCACTTTCTT    3770

Qy   1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAAGTGCCCTACTAT     1980
          |||||||||||||||||||
```

```
Db   3771  TCCACGTCGTCGTCCAGGTGCCTGGGAGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 3830

Jy   1981  ATCAACCACGAGACTCAAACA 2001
           ||||||||||||||||||||
Db   3831  ATCAACCACGAGACTCAAACA 3851

RESULT 8
AAD37230
AD37230  standard; DNA; 4182 BP.
X
C  AAD37230;
X
T  21-AUG-2002  (first entry)
X
E  Human dystrophin minigene delta4173.
X
W  Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W  adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W  Becker muscular dystrophy; ds.
X
S  Homo sapiens.
X
N  WO20018369S-A2.
X
D  08-NOV-2001.
X
F  27-APR-2001; 2001WO-US013677.
R  28-APR-2000; 2000US-0200777P.
X
A  (XIAO/) XIAO X.
I  Xiao X;
X
R  WPI; 2002-049342/06.
X
T  New dystrophin minigene for treating Duchenne or Becker muscular
T  dystrophy comprises an N-terminal domain or modified N-terminal domain,
T  rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
T  gene.
X
S  Example 1; Page 43-44; 71pp; English.
X
C  The present invention relates to an isolated nucleotide sequence encoding
C  a dystrophin minigene. The minigene comprises N-terminal or modified N-
C  terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
C  domains and cysteine-rich domains of dystrophin or utrophin genes. The
C  invention also relates to a recombinant adeno-associated virus (AAV)
C  comprising dystrophin minigene operably linked to an expression control
C  element. The dystrophin minigene in operable linkage with an expression
C  control element, in a recombinant adeno-associated virus or retrovirus is
C  useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
C  dystrophy (BMD) in a mammalian subject. The present sequence is human
C  dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
C  hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
C  hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
X
Q  Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;

Query Match     83.1%;  Score 1662.8;  DB 6;  Length 4182;
Best Local Similarity  88.3%;  Pred. No. 0;
Matches 1929;  Conservative  0;  Mismatches  72;  Indels 183;  Gaps  4;
```

```
Db  1120  GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGGATTGACA 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   181  GCCCATCAGGGCACCGGCTTGGTTGGTAATATTCTACAATTGGGAAGTAGCTGATTGGAACAGGA 240
          |||||||||||||||||||||||||   ||| |||||||||||||||||||||||||||
Db  1180  GCCCATCAGGGGCCGGGTTGGTTAATATTCTACAATTGGGAAGTAGCTGATTGGAACAGGA 1239

Qy   241  AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1240  AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1299

Qy   301  TGGGAATGCCTCAGGGTAGCTAGCATGGCATGGAAAAACAAAGCAATTACATAGAGTTTAATG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1300  TGGGAATGCCTCAGGGTAGCTAGCATGGCATGGAAAAACAAAGCAATTACATAGAGTTTAATG 1359

Qy   361  GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1360  GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 1419

Qy   421  ACAAGGAAAATGGAGGAAGAGCCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1420  ACAAGGAAAATGGAGGAAGAGCCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTA 1479

Qy   481  CAACAACATAGGTGCTTCAAGAGATCAGAACAAGAGCAAGTCAGGGTCAATTCTCTCC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1480  CAACAACATAGGTGCTTCAAGAGATCAGAACAAGAGCAAGTCAGGGTCAATTCTCTCC 1539

Qy   541  ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGAATCACGCAACTGCTGTTTGGAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1540  ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGAATCACGCAACTGCTGTTTGGAA 1599

Qy   601  GAACAACTTAAGGTATTGGGAGATCGATGGCAAACATCTGTAGGACAGAAGACCGC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1600  GAACAACTTAAGGTATTGGGAGATCGATGGCAAACATCTGTAGGATGAGCAGAAGACCGC 1659

Qy   661  TGGGTTCTTTTACAAGACCAGCCTGAC---------CTAGCTCCTGGACTGACCACT 708
          |||||||||||||||||||||||||||         ||||||||||||||||||||||
Db  1660  TGGGTTCTTTTACAAGACACATCCTTCTCAAATGGCAAGTTCTTACTGAAGAACAGTGCCTT 1719

Qy   709  ATTGGAGGCCTCCTCCTACTCCAGACTGTTACTCTGGTGACCA-------ACCTGTGGTT 759
          ||||||||||||||||||||||||||||||||||||||||||       |||||||||
Db  1720  TTTAGTGCATGGCTTTCAGAAAAAGAAGAATGGCAGTGACACAAGAATTCACACAACTGGCTTT 1779

Qy   760  ACTAAGGAAACTGCCATCTC------------ 779
          ||||||||||||||||||||
Db  1780  AAAGATCAAAATGAAAATGTTATCAAGTCTTCAAAACTGGCCGTTTAAAAGCGGATCTA 1839

Qy   780  ---------------------CAAACTAGAAAATGCCATCTTCC 801
                               ||||||||||||||||||||
Db  1840  GAAAAGAAAAAGCAATCCATGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACA 1899

Qy   802  TTGATGTTGGAG---------------------- 813
          |||||||||||
Db  1900  CTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGCTGGATAACTTTGCCGGTGT 1959

Qy   814  ------------------GTACCTACTCATAGATTACTGCAA 837
                            ||||||||||||||||||||||
Db  1960  TGGGATAATTTAGTCCAAACTTGAAAGAGTACAGCACAGACTCATATAGATTACTGCAA 2019

Qy   838  CAGTTCCCCTGGACCTGGAAAGTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCC 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2020  CAGTTCCCCTGGACCTGGAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCC 2079

Qy   898  AATGTCCTACAGGATGTCACCGTAAGGAAAGGCTCCTAGAAGAAAGCTTTAGAAGAA 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2080  AATGTCCTACAGGATGTCACCGTAAGGAAAGGCTCCTAGAAGAAAGCTTTAGAAGAA 2139

Qy   958  GAGCTGATGAAACAATGGCAAGACCTCCAGGTGAAATTGAAGCTCACACAGATGTTTAT 1017
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2140  GAGCTGATGAAACAATGGCAAGACCTCCAGGTGAAATTGAAGCTCACACAGATGTTTAT 2199

Qy  1018  CACACCTGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGAAGTTCCGATGATGCA 1077
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy     1  GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGGTTATCAAACAGCTTTAGAAGAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1000  GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGGTTATCAAACAGCTTTAGAAGAA 1059

Qy    61  GTATTATCGGTGCTTCTTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1060  GTATTATCGGTGCTTCTTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1119

Qy   121  GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGGTACATGATGGATTGACA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 57-59; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;

Query Match             83.1%;  Score 1662.8;  DB 6;  Length 5149;
Best Local Similarity   88.3%;  Pred. No. 0;
Matches 1929;  Conservative   0;  Mismatches  72;  Indels 183;  Gaps   4;

Qy          1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1757 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1816

Qy         61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1876

Qy        121 GATGTGGAAGTGTGAAAGACCAGTTTCATCTCATGAGGGGTACATGATGATGGATTTGACA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1877 GATGTGGAAGTGTGAAAGACCAGTTTCATCTCATGAGGGGTACATGATGATGGATTTGACA 1936

Qy        181 GCCCATCAGGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1937 GCCCATCAGGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996

Qy        241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGGAGATTCTAAT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1997 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGGAGATTCTAAT 2056

Qy        301 TGGGAATGCCTCAGGGTAGCTCAGGTTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 360
```

```
Db       2200 CACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 2259

Qy       1078 GTCCTGTTACAAAGACGTTTGGATAACATGAAGTGAGTGAACTTCGGAAAAAG 1137
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2260 GTCCTGTTACAAAGACGTTTGGATAACATGAAGTGAGTGAACTTCGGAAAAAG 2319

Qy       1138 TCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2320 TCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTT 2379

Qy       1198 TCTCTGCGAGGAACTTCTGTGTGGCTACAGTGAAAGATGATGAATTAAGCCGGCAGGCA 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2380 TCTCTGCGAGGAACTTCTGTGTGGCTACAGTGAAAGATGATGAATTAAGCCGGCAGGCA 2439

Qy       1258 CCTATTGGAGGGCGACTTCCAGCAGTTCAGAGCAGCAGAACGATGTACATAGGGCCTTCAAG 1317
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2440 CCTATTGGAGGGCGACTTCCAGCAGTTCAGAGCAGCAGAACGATGTACATAGGGCCTTCAAG 2499

Qy       1318 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT 1377
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2500 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTT 2559

Qy       1378 CTGACAGAGCAGCAGCCTTTGGAAGGACTAGAGAAAACTAGAACTCTACCAGGAGCCCAGAGAGCTGCCT 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2560 CTGACAGAGCAGCAGCCTTTGGAAGGACTAGAGAAAACTAGAACTCTACCAGGAGCCCAGAGAGCTGCCT 2619

Qy       1438 CCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAAT 1497
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2620 CCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAAT 2679

Qy       1498 ACTGAGTGGGAAAAATTGAACCTGACTCCCGCTGACTGGCAGAGAAAAATAGATGAGACC 1557
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2680 ACTGAGTGGGAAAAATTGAACCTGACTCCCGCTGACTGGCAGAGAAAAATAGATGAGACC 2739

Qy       1558 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTGCGCCAA 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2740 CTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTGCGCCAA 2799

Qy       1618 GCTGAGGTGATCAAGGGATCCTCGCAGCCCGTGGGCCATCCCATTCCTCATTGACTCTCCAA 1677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2800 GCTGAGGTGATCAAGGGATCCTCGCAGCCCGTGGGCCATCCCATTCCTCATTGACTCTCCAA 2859

Qy       1678 GATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTGCCGCCTCTGAAAGAGAACGTG 1737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2860 GATCACCTCGAGAAAAGTCAAGGCACTTCGAGGAGAAATTGCCGCCTCTGAAAGAGAACGTG 2919

Qy       1738 AGCAGCGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTTCACCGTAT 1797
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2920 AGCAGCGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTTCACCGTAT 2979

Qy       1798 AACCTCAGCACTCTGGAGAGCTCGGAGAGACCTGAACACCAGATGAGAAGCTTCTGCAGGTTCTGCGCAGTTT 1857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2980 AACCTCAGCACTCTGGAGAGCTCGGAGAGACCTGAACACCAGATGAGAAGCTTCTGCAGGTTCTGCGCAGTTT 3039

Qy       1858 GACCGAGTCAGGCAGCTGCATGAGACCCACAGGGACTTTGGTCCAGCATCTCAGCACTTT 1917
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3040 GACCGAGTCAGGCAGCTGCATGAGACCCACAGGGACTTTGGTCCAGCATCTCAGCACTTT 3099

Qy       1918 CTTTCCACGTCGTCCAGGGTCCCTGGGAGAGCCATCTCCCAAACAAAGTGCCCTAC 1977
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3100 CTTTCCACGTCGTCCAGGGTCCCTGGGAGAGCCATCTCCCAAACAAAGTGCCCTAC 3159

Qy       1978 TATATCAACCACGAGACTCAAACA 2001
              ||||||||||||||||||||||||
Db       3160 TATATCAACCACGAGACTCAAACA 3183
```

```
RESULT 9
AAD37255
ID   AAD37255  standard; DNA; 5149 BP.
XX
KX   AAD37255;
XX
KX   AAD37255;
```

```
Qy    2057 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 2116
Db    3137 TCTCTGCAGGAACTTCTGGTGTGCACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 3196

Qy     361 GATCTCCAGAATCAGAAACTCAGAAGATCAGAAGAGTTGAATGACTGGCTAACAAAAACACAGAAGAAAGA 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1258 CCTATTGGAGCGACTTCAGCAGTTCAGAAGCAGCAGATGTACATAGGGCCTTCAAG 1317

Qy    2117 GATCTCCAGAATCAGAAACTCAGAAGATGAATGACTGGCTAACAAAAACACAGAAGAAAGA 2176
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3197 CCTATTGGAGCGAGCGACTTCAGCAGTTCAGAGAGCAGAACGATGTACATAGGGCCTTCAAG 3256

Qy     421 ACAAGGAAAATGGAGGAAGAGCCTTCTTGGACCTGATCTTGAAGAACCTAAAACGCCAAGTA 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1318 AGGGAATTGAAAACTAAAGAACTGTAATCATCAGAAGAACTTCATCAGAGAGCTGTACGAATATTT 1377

Qy    2177 ACAAGGAAAATGGAGGAAGAGCCTTCTTGGACCTGATCTTGAAGAACCTAAAACGCCAAGTA 2236
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3257 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTAGTCTTGAGACTGTACGATATATT 3316

Qy     481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1378 CTGACAGAGCAGCGCTTTGGAAGGACTAGAGAAACTTCACCAGGACGCCCAGAGAGCTGCCT 1437

Qy    2237 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 2296
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3317 CTGACAGAGCAGCGCTTTGGAAGGACTAGAGAAACTTCACCAGGACGCCCAGAGAGCTGCCT 3376

Qy     541 ACTCACATGGTGGTGTAGTTGATGAATTCAGGTGGAGATCACCGCAACTGCTGCTTTGGAA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1438 CCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCACGAAAGTGCAGGCTGAGGAGGTCAAT 1497

Qy    2297 ACTCACATGGTGGTGTAGTTGATGAATTCAGGTGGAGATCACCGCAACTGCTGCTTTGGAA 2356
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3377 CCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGTCAAT 3436

Qy     601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCGTAGTAGTGGACAGAGACCGC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1498 ACTGAGTGGGAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACC 1557

Qy    2357 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCGTAGTAGTGGACAGAGAAGACCGC 2416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3437 ACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 3496

Qy     661 TGGGTTCTTTTACAAGACCAGCCTGAC------CTAGCTCCTGGACTGACCACT 708
           |||||||||||||||||||||||||||      |||||||||||||||||||||
Db    1558 CTTGAAAGACTTCCAGGAACTTCAAGAGACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAA 1617

Qy    2417 TGGGTTCTTTTACAAGACATCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTT 2476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3497 CTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAA 3556

Qy     709 ATTGGAGCCTCCACTCGACTGTTACTCTGGTGACACA--------ACCTGTGTT 759
           |||||||||||||||||||||||||||||||||||||        |||||||||
Db    1618 GCTGAGGTGATCAAGGGACTCCTGGCAGCCGCCGGTGGGCGATCTCCTCCATGACTTCTCTCCAA 1677

Qy    2477 TTTAGTGCATGGCTTTCAGAAAAAGAAGAGATGCGAGTGAACAAGATTCACCACACTGGCTTT 2536
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3557 GCTGAGGTGATCAAGGGACTCCTGGCAGCCGCCGGTGGGCGATCTCCTCATTGACTCTCTCCAA 3616

Qy     760 ACTAAGGAAACTGCCATCTC------------------- 779
           |||||||||||||||                        
Db    1678 GATCACCTCGAGAAAGTCAAGGCACTTCGAGGACACTTCGCGCCTCTGAAAAGAGAAACGTG 1737

Qy    2537 AAAGATCAAAATGAAAATGTTATCAAGTCTTCAAAAAACTGCCGTTTTAAAAGCGGATCTA 2596
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3617 GATCACCTCGAGAAAGTCAAGGCACTTCGAGGACACTTCGCGCCTCTGAAAAGAGAAACGTG 3676

Qy     780 ------------------------------CAAACTAGAAATGCCATCTCC 801
                                         ||||||||||||||||||||
Db    1738 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTAT 1797

Qy    2597 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAACAAGATCTTCTTCTTCAACA 2656
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3677 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTAT 3736

Qy     802 TTGATGTTGGAG------- 813
           ||||||||||          
Db    1798 AACCTTGCAGCACTCTGAAGACTCTGAAGACCAGATGGAAGCTTCTGCAGTGGCCGTCGAG 1857

Qy    2657 CTGAAGAATAGTCAGTGACCAGAAGACGGAAGCATGCTGGATAACTTGCCCGGTGT 2716
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3737 AACCTTGCAGCACTCTGAAGACTCTGAAGACCAGATGGAAGCTTCTGCAGTGGCCGTCGAG 3796

Qy     814 ------------------GTACCTACTCATAGATTACTGCAA 837
                             ||||||||||||||||||||||
Db    1858 GACCGAGTCAGGCAGCTGCATGAGAGCTGCATGAGAGCCCACAGGGACGACTTTGGTCCAGCATCTCAGCACTTT 1917

Qy    2717 TGGGATAATTTAGTCCAAAAACTTGAAAAGAGTACACCAGACTCAATAGATTACTGCAA 2776
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3797 GACCGAGTCAGGCAGCTGCATGAGAGCTGCATGAGAGCCCACAGGGACGACTTTGGTCCAGCATCTCAGCACTTT 3856

Qy     838 CAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCC 897
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1918 CTTTCCACGTCTGTCCAGGTGCCCTGGGAGGAGAGCCCATCTCGCCAAACAAAGTGCCCTAC 1977

Qy    2777 CAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCC 2836
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3857 CTTTCCACGTCTGTCCAGGTGCCCTGGGAGGAGAGCAGCCATCTCGCCAAACAAAGTGCCCTAC 3916

Qy     898 AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCCTCCTAGAAGACTTCCAAGGGAGTAAAA 957
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1978 TATATCAACCACCACGACTCAAACA 2001

Qy    2837 AATGTCCTACAGGATGCTACCCGTAAGGAAAGGCCTCCTAGAAGACTTCCAAGGGAGTAAAA 2896
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3917 TATATCAACCACCACGACTCAAACA 3940

Qy     958 GAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 1017
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    2897 GAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 2956

Qy    1018 CACAACCTTGGATGAAAACAGCCAAAAAATCCTCGGATCCCTGGAAGGTTCCGATGATGCA 1077
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    2957 CACAACCTTGGATGAAAACAGCCAAAAAATCCTCGGATCCCTGGAAGGTTCCGATGATGCA 3016

Qy    1078 GTCCTGTTACAAAGACGTTTGGATAACAGCAACTTCAAGTGGAGTGAACTTCGGAAAAAG 1137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    3017 GTCCTGTTACAAAGACGTTTGGATAACAGCAACTTCAAGTGGAGTGAACTTCGGAAAAAG 3076

Qy    1138 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTT 1197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    3077 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTT 3136

Qy    1198 TCTCTGCAGGAACTTCTGGTGTGCACAGCTGAAAGATGATGAATTAAGCCGGCAGGCA 1257
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 10
ABK81999
ID   ABK81999 standard; DNA; 5462 BP.
XX
AC   ABK81999;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaR2-R21+H3.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO20229056-A2.
XX

PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 14; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 U; 0 Other;

Query Match   64.1%; Score 1283; DB 6; Length 5462;
Best Local Similarity   99.3%; Pred. No. 0;
Matches 1302;   Conservative   0; Mismatches   0; Indels   9; Gaps   1;

Qy   691 GCTCCTGGACTGACCACTATTGGAGCGCCTCCACTCGACTGTTACTCTGGTGACACAA 750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 GCTCCTGGACTGACCACTATTGGAGCCTTCCACTCGACTGTTACTCTGGTGACACAA 1606

Qy   751 CCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAACTAGAAAAAGTTCTTGGT 810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 CCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAACTAGAAAAAGTTCTTGGT 1666

Qy   811 GAGGTACCTACTCATAGATTACTGCAACAGTCCCCCTGGACCTGGAAAAAGTTCTTGCC 870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1667 GAG-------CATAGATTACTGCAACAGTCCCCCTGGACCTGGAAAAAGTTCTTGCC 1717

Qy   871 TGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAG 930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1718 TGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAG 1777

Qy   931 CTCCTAGAAGACTTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGT 990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1778 CTCCTAGAAGACTTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGT 1837

Qy   991 GAAATTGAAGCTCACACAGATGTTTATCACAACTGATGAAAACAGCCAAAAAATCCTG 1050
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1838 GAAATTGAAGCTCACACAGATGTTTATCACAACTGATGAAAACAGCCAAAAAATCCTG 1897

Qy   1051 AGATCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAAC 1110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1898 AGATCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAAC 1957

Qy   1111 TTCAAGTGAGTGAACTTCGGAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGT 1170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1958 TTCAAGTGAGTGAACTTCGGAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGT 2017

Qy   1171 TCTGACCAGTGGAAGCGTCTGCACCTTTCTGCACCTTTCTGCGAGGAACTTCGGTGTGCTACAGCTG 1230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2018 TCTGACCAGTGGAGAGCGTCTGCACCTTTCTGCACCTTTCTGCGAGGAACTTCGGTGTGCTACAGCTG 2077

Qy   1231 AAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAG 1290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2078 AAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGCACTTTCCAGCAGTTCAGAAG 2137

Qy   1291 CAGAACGATGTACATAGGCCTTCAGCAGGGAATTGAAAACTAAAGAACCTGTAATCATG 1350
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2138 CAGAACGATGTACATAGGCCTTCAGCAGGGAATTGAAAACTAAAGAACCTGTAATCATG 2197

Qy   1351 AGTACTCTTGAGACTGTACGAATATTTCGACAGGAGCAGCCTTGGAAGGACTAGAGAAA 1410
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2198 AGTACTCTTGAGACTGTACGAATATTTCGACAGGAGCAGCCTTTGGAAGGACTAGAGAAA 2257

Qy   1411 CTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAGAGGCCCAGAATGTCACTCGGCTT 1470
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2258 CTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAGAGGCCCAGAATGTCACTCGGCTT 2317

Qy   1471 CTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCT 1530
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2318 CTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCT 2377

Qy   1531 GACTGGCAGAGAAAATATGATGAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGCCACG 1590
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2378 GACTGGCAGAGAGAAAAATGATGAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGCCACG 2437

Qy   1591 GATGAGCTGGACTTGCAAGCTGGGCGTGAGGTGATCAAGGGTGATCAAGGAAAGTCTGCAGCCCGTG 1650
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2438 GATGAGCTGGACTTGCAAGCTGGGCGTGAGGTGATCAAGGGTGATCAAGGAAAGTCTGGCAGCCCGTG 2497

Qy   1651 GGGCATCTCCTCATTGACTCCTCCAAGATCACCTGAGAAGTCAAGGCACTTGGAGGA 1710
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2498 GGGCATCTCCTCATTGACTCCTCCAAGATCACCTGAGAAGTCAAGGCACTTGGAGGA 2557

Qy   1711 GAAATTGCCCTGAAAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACC 1770
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2558 GAAATTGCCCTGAAAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACC 2617

Qy   1771 ACTTTGGGCATTCAGCTTCACCGTATAACCTCGAAGACTCTGGAAGACCTGAACACCAGA 1830
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2618 ACTTTGGGCATTCAGCTTCACCGTATAACCTCGAAGACTCTGGAAGACCTGAACACCAGA 2677

Qy   1831 TGGAAGCTTCTGCAGGTGGCCGTCGAGGAGAGCCAGTCGAGAGCCCACAGG 1890
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2678 TGGAAGCTTCTGCAGGTGGCCGTCGAGGAGACCGAGTCGAGAGCCGCCACAGG 2737

Qy   1891 GACTTTGGTCCAGCATCCAGCACTTCTTCCACGTCGTGTCCAGGGTCCTGGGAGAGA 1950
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2738 GACTTTGGTCCAGCATCCAGCACTTCTTCCACGTCGTGTCCAGGGTCCTGGGAGAGA 2797

Qy   1951 GCCATCTCGCCAAACAAAGTGCCCAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2798 GCCATCTCGCCAAACAAAGTGCCCAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2848

RESULT 11
AAD06794
ID   AAD06794 standard; DNA; 5952 BP.
XX
AC   AAD06794;
XX
DT   06-AUG-2001 (first entry)
XX
DE   Human dystrophin gene (Becker form).
XX
KW   Human; dystrophin; extein; intein; trans-splicing; gene therapy;
KW   Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   misc_recomb   2847..2848
FT   /tag= a
FT   /label= S4_junction_site
FT   /note="Dystrophin gene is split at this site and
FT   attached to an intein sequence. The resulting fragment is
FT   used to produce plasmid pSD4"

FT misc_recomb   2952..2953
FT   /*tag= b
FT   /label= S3_junction_site
FT   /note= "Dystrophin gene is split at this site and
FT   attached to an intein sequence. The resulting fragment is
FT   used to produce plasmid pSD3"
FT misc_recomb   3198..3199
FT   /*tag= c
FT   /label= S2_junction_site
FT   /note= "Dystrophin gene is split at this site and
FT   attached to an intein sequence. The resulting fragment is
FT   used to produce plasmid pSD2"
FT misc_recomb   3300..3301
FT   /*tag= d
FT   /label= S1_junction_site
FT   /note= "Dystrophin gene is split at this site and
FT   attached to an intein sequence. The resulting fragment is
FT   used to produce plasmid pSD1"
XN WO200129243-A1.
XD 26-APR-2001.
XX 13-OCT-2000; 2000WO-CA001216.
XF 15-OCT-1999; 99US-0159868P.
XR
XA (UYDA-) UNIV DALHOUSIE.
XA (UYPI-) UNIV PITTSBURGH.
XA Paul XL, Xiao X;
XR WPI; 2001-367297/38.
FT
FT Use of spontaneous or automatic protein splicing to join two or more
FT peptides at junction site involves expressing extein peptides having co-
FT reacting portions of split intein attached to them, so that peptides
XX splice.
XS Example 1; Fig 2; 81pp; English.
CC The invention relates to a method directed to the use of spontaneous or
CC automatic protein trans-splicing to join two or more peptides at junction
CC site. This method involves expressing extein peptides having co-reacting
CC portions of split intein attached to them, so that peptides will splice
CC automatically under suitable conditions. The invention also provides
CC methods for circumventing virion packaging size limitations in
CC recombinant virus particle, by splitting a coding region for a protein to
CC be delivered into two or more extein genes, which are packed into
CC separate virus particles and are co-delivered in a target cell for the
CC expression and for subsequent trans-splicing to form the complete
CC protein. In particular, the method is used for trans-splicing human
CC dystrophin and in gene therapies of recombinant adeno-associated virus
CC (AAV) particles that encode trans-spliced dystrophin, for treating
CC diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular
CC dystrophy (BMD). The present sequence is human dystrophin gene (Becker
XX form)

Q Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 U; 0 Other;
   Query Match   59.1%;   Score 1182.6;   DB 5;   Length 5952;
   Best Local Similarity   99.7%;   Pred. No. 0;
   Matches 1185;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;

Y   813 GGTACCTACTCATCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 872
          |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2946 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAAGTTTCTTGCCTG 3005
Y   873 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3006 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 3065
Y   933 CCTAGAAGACTCCAAGGAGGAGTAAAGAAGCTGATGAAACAATGGCAAGACTTCCAAGGTGA 992

Db   3066 CCTAGAAGACTCCAAGGAGGAGTAAAGAAGCTGATGAAACAATGGCAAGACTTCCAAGGTGA 3125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   993 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAACAGCCAAAAAATCCTGAG 1052
Db   3126 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAACAGCCAAAAAATCCTGAG 3185
Qy   1053 ATCCCTGGAAGGTTCCGATGATGCAGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT 1112
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3186 ATCCCTGGAAGGTTCCGATGATGCAGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT 3245
Qy   1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTTCAACATTAGGTCCCATTTGGAAGCCAGTTC 1172
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   3246 CAAGTGGAGTGAACTTCGGAAAAAGTCTCCAACATTAGGTCCCATTTGGAAGCCAGTTC 3305
Qy   1173 TGACCAGTGGAAGCGTCTGCAGACTTCTCGGTGTGGCTACGCTGAA 1232
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3306 TGACCAGTGGAAGCGTCTGCAGACTTCTCGGTGTGGCTACGCTGAA 3365
Qy   1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3366 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 3425
Qy   1293 GAACGATGTACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3426 GAACGATGTACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAACCTGTAATCATGAG 3485
Qy   1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGGAGCCTTGGAAGGACTAGAGAAACT 1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3486 TACTCTTGAGACTGTACGAATATTTCTGACAGAGGAGCCTTGGAAGGACTAGAGAAACT 3545
Qy   1413 CTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGGAGCCCAGAATGTCACTCGGCTTCT 1472
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3546 CTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGGAGCCCAGAATGTCACTCGGCTTCT 3605
Qy   1473 ACGAAAGCAGGCAGCGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 1532
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3606 ACGAAAGCAGGCAGCGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 3665
Qy   1533 CTGGCAGAGAAAATAGATGAGACCCTGAAAGACTCCAAGGACTTCAAGAGGCCACGCGA 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   3666 CTGGCAGAGAAAATAGATGAGACCCTGAAAGACTCCAAGGACTTCAAGAGGCCACACGGA 3725
Qy   1593 TGAGCTGGACCTCAAGCTGCGGGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3726 TGAGCTGGACCTCAAGCTGCGGGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG 3785
Qy   1653 CGATCTCCTCATTGACTCCTCCCAAGATCACCTCCGAGAAAGTCAAGGCCACTTCGGAGGA 1712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3786 CGATCTCCTCATTGACTCCTCCCAAGATCACCTCCGAGAAAGTCAAGGCCACTTCGGAGGA 3845
Qy   1713 AATTGCCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3846 AATTGCCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3905
Qy   1773 TTTGGGCATTCAGCTCTCACCGGTATAACCTCGAAGACCTCGAAGACACCAGATG 1832
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3906 TTTGGGCATTCAGCTCTCACCGGTATAACCTCGAAGACCTCGAAGACACCAGATG 3965
Qy   1833 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGCAGCTGCATGAGCCCACAGGGA 1892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3966 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGCAGCTGCATGAGCCCACAGGGA 4025
Qy   1893 CTTTGGTCCAGCATCCAGCACTTCTTCCACGTCGTGTCCAGGTCCTGGGAGAGC 1952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   4026 CTTTGGTCCAGCATCCAGCACTTCTTCCACGTCGTGTCCAGGTCCTGGGAGAGAGC 4085
Qy   1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4086 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 4134

Y RESULT 12

```
ABK82000
ID   ABK82000 standard; DNA; 8689 BP.
XX
AC   ABK82000;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaH2-R19.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200229056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 15; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 U; 0 Other;

Query Match       59.1%;    Score 1182.6;    DB 6;    Length 8689;
Best Local Similarity   99.7%;    Pred. No. 0;
Matches 1185;    Conservative 0;    Mismatches 4;    Indels 0;    Gaps 0;

Qy  813 GGTACCTACTACATCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG  872
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2992 GGAAGAAACTACATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 3051

Qy  873 GCTTACAGAGAGCTGAAACAACACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT  932
        ||||||||||||||||||||||||||||||||
Db 3052 GCTTACAGAGAGCTGAAACAACACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 3111

Qy  933 CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA  992
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3112 CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 3171

Qy  993 AATTGGAAGCTTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAAAATCCTGAG 1052
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3172 AATTGGAAGCTTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAAAATCCTGAG 3231

Qy 1053 ATCCCTGGAAGTTCCGATGATGCAGTCAGTTCCTGTTACAAAGACGTTTGGATAACATGAACTT 1112
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3232 ATCCCTGGAAGTTCCGATGATGCAGTCAGTTCCTGTTACAAAGACGTTTGGATAACATGAACTT 3291

Qy 1113 CAAGTGGAGTGAACTTCGGGAAAAAGTCTCTCAACATTAGGTCCCCATTTGGAAGCCAGTTC 1172
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3292 CAAGTGGAGTGAACTTCGGGAAAAAGTCTCTCAACATTAGGTCCCCATTTGGAAGCCAGTTGAA 3351

Qy 1173 TGACCAGTGGAAGCGTCTGCACCTTCTGCGAGGACACCATTGGTGTGTGGCTACAGCTGAA 1232
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3352 TGACCAGTGGAAGCGTCTGCACCTTCTGCGAGGACACCATTGGTGTGTGGCTACAGCTGAA 3411

Qy 1233 AGATGATGAATTAAGCCGGCAGGGCACCCATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 1292
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3412 AGATGATGAATTAAGCCGGCAGGGCACCCATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 3471

Qy 1293 GAACGATGTACATAGGGCCTTCAAGAGGAGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3472 GAACCATGTACATAGGGCCTTCAAGAGGAGGAATTGAAAACTAAAGAACCTGTAATCATGAG 3531

Qy 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACT 1412
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3532 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACT 3591

Qy 1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCT 1472
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3592 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCT 3651

Qy 1473 ACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATGAACCTGCACTCCGCTGA 1532
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3652 ACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 3711

Qy 1533 CTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCAAGAGACTTCAAGAGAGCCACCGGA 1592
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3712 CTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCAAGAGACTTCAAGAGAGCCACCGGA 3771

Qy 1593 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGTGGG 1652
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3772 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGTGGG 3831

Qy 1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGAGAGA 1712
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3832 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGAGAGA 3891

Qy 1713 AATTGCGCCTCTGAAAGAAGACGTGAGCGACAGTGAGCCGCACGTCAATGACCTTGCTCGCCCAGCTTACCAC 1772
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3892 AATTGCGCCTCTGAAAGAAGACGTGAGCGACAGTGAGCCGCACGTCAATGACCTTGCTCGCCCAGCTTACCAC 3951

Qy 1773 TTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 1832
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3952 TTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 4011

Qy 1833 GAAGGTTCTGCAGGTGGGCCGTCGAGGACCGAGCCAGCTGCATGAGAAGCCCACAGGGA 1892
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4012 GAAGGTTCTGCAGGTGGGCCGTCGAGGACCGAGCCAGCTGCATGAGAAGCCCACAGGGA 4071

Qy 1893 CTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCTGTCAGGGTCCTGGGAGAGC 1952
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4072 CTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCTGTCAGGGTCCTGGGAGAGC 4131
```

```
RESULT 13
AAD37229
ID   AAD37229 standard; DNA; 11058 BP.
XX
AC   AAD37229;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin protein coding sequence.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
```

```
Qy  1293  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8532  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  8591

Qy  1353  TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTGAAGACTAGAGAAACT  1412
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8592  TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGACTAGAGAAACT  8651

Qy  1413  CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCT  1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8652  CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCT  8711

Qy  1473  ACGAAAGCAGGCTCAATACTGAGTGGGAAAATTGAACCTGCACTCCGCTGA  1532
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  8712  ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA  8771

Qy  1533  CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAGGAGACTTCAGAGGGCCACGA  1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8772  CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAGGAGACTTCAGAGGGCCACGGA  8831

Qy  1593  TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG  1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8832  TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG  8891

Qy  1653  CGATCTCCTCATTGACTCTCTCCAAGATCACTCTGAGAAAGTCAAGGCACTTCGAGGAGA  1712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8892  CGATCTCCTCATTGACTCTCTCCAAGATCACTCTGAGAAAGTCAAGGCACTTCGAGGAGA  8951

Qy  1713  AATTGGCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCGAGCTTACCAC  1772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8952  AATTGGCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCGAGCTTACCAC  9011

Qy  1773  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCAGATG  1832
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9012  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGGAACACCAGATG  9071

Qy  1833  GAAGCTTCTGCAGGTGGCCGTGCCGAGGACGGACCCAGCGAGTCAGCGAGCCTGCATGAGAGCCCACAGGGA  1892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9072  GAAGCTTCTGCAGGTGGCCGTGCCGAGGACGGACCCAGCGAGTCAGCGAGCCTGCATGAGAGCCCACAGGGA  9131

Qy  1893  CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCGTCCAGGTCCTGGGAGAGAGC  1952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9132  CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCGTCCAGGTCCTGGGAGAGAGC  9191

Qy  1953  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA  2001
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  9192  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA  9240
```

```
X T   New dystrophin minigene for treating Duchenne or Becker muscular
X T   dystrophy comprises an N-terminal domain or modified N-terminal domain,
X T   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
X T   gene.
X S   Example 1; Page 40-43; 71pp; English.
X C   The present invention relates to an isolated nucleotide sequence encoding
X C   a dystrophin minigene. The minigene comprises N-terminal or modified N-
X C   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
X C   domains and cysteine-rich domains of dystrophin or utrophin genes. The
X C   invention also relates to a recombinant adeno-associated virus (AAV)
X C   comprising dystrophin minigene operably linked to an expression control
X C   element. The dystrophin minigene in operable linkage with an expression
X C   control element, in a recombinant adeno-associated virus or retrovirus is
X C   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
X C   dystrophy (BMD) in a mammalian subject. The present sequence is human
X C   dystrophin protein coding sequence
X Q   Sequence 11058 BP; 3686 A; 2292 C; 2621 G; 2459 T; 0 U; 0 Other;
```

```
Query Match     59.1%;  Score 1182.6;  DB 6;  Length 11058;
Best Local Similarity   99.7%;  Pred. No. 0;
Matches 1185;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    0;
```

```
Qy   813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTG  872
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8052  GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTG  8111

Qy   873  GCTTACAGAGCTGAAACAACTGCCAATGTCCTACCAGGATGCTACCCGTGAAGGAAAGGCT  932
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8112  GCTTACAGAGCTGAAACAACTGCCAATGTCCTACCAGGATGCTACCCGTGAAGGAAAGGCT  8171

Qy   933  CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGAACCTCCAAGGTGA  992
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8172  CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGAACCTCCAAGGTGA  8231

Qy   993  AATTGAAGCTCACACAGATGTTATCACAACCTGATGAAACAGCCAAAAATCCTGAG  1052
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8232  AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAACAGCCAAAAATCCTGAG  8291

Qy  1053  ATCCCTGGAAGTTCCGATGATGCAGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACT  1112
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8292  ATCCCTGGAAGTTCCGATGATGCAGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT  8351

Qy  1113  CAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  1172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8352  CAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  8411

Qy  1173  TGACCAGTGGAAGCGTCTGCACCTTCGCGAGCCACCTTCGTGTGCTACAGCTGAA  1232
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8412  TGACCAGTGGGAAGCGTCTGCACCTTCTCGGTGTGCTACAGCTGAA  8471

Qy  1233  AGATGATGAATTAAGCCGGCACCGCAGCCAGGCACCTATTGGAGGAGCTCAGAAGCA  1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8472  AGATGATGAATTAAGCCGGCACCGCAGCCAGGCACCTATTGGAGGAGCTCAGAAGCA  8531
```

```
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Example 2; Fig 23; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for the therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a human dystrophin polynucleotide
CC   sequence used in the creation of the mini-dystrophin peptides of the
CC   invention
XX
SQ   Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 U; 0 Other;

Query Match             59.1%;  Score 1182.6;  DB 6;  Length 11241;
Best Local Similarity   99.7%;  Pred. No. 0;
Matches 1185;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy    813 GGTACCTACTCATAGAATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTG  872
          ||||||                |||||||||||||||||||||||||||||||||||||
Db   8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTG 8319

Qy    873 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT  932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8320 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 8379

Qy    933 CCTAGAGAGACTCCAAGGAGTAAAAGAGAGCTGATGAAACAATGGCAAGACCTCCAAGTGA  992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8380 CCTAGAGAGACTCCAAGGAGTAAAAGAGAGCTGATGAAACAATGGCAAGACCTCCAAGTGA 8439

Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCCTGAG 1052
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCCTGAG 8499

Qy   1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAGACGTTTGGATAACATGAACTT 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8500 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAGACGTTTGGATAACATGAACTT 8559

Qy   1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAGCCAGTTC 1172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8560 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAGCCAGTTC 8619

Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGACTTCTCTGTGTGTGGCTACAGCTGAA 1232
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8620 TGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGACTTCTCTGTGTGTGGCTACAGCTGAA 8679

Qy   1233 AGATGATGAATTAAGCGCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8680 AGATGATGAATTAAGCGCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 8739

Qy   1293 GAACGATGTACATAGGGCCTTCAGAGAGGAATTGAAACTAAAGAACCTGTAATCATGAG 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8740 GAACGATGTACATAGGGCCTTCAGAGAGGAATTGAAACTAAAGAACCTGTAATCATGAG 8799

Qy   1353 TACTCCTTGAGACTGTACGAATATTTCTGACAGAGGCCTTTGGAAGGAGACTAGAGAAACT 1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8800 TACTCCTTGAGACTGTACGAATATTTCTGACAGAGGCCTTTGGAAGGACTAGAGAAACT 8859

Qy   1413 CTACCAGGAGCGCCCTGCCCTGAGAGAGTCGCCCAGAGAGCCAGAATGTCACTCGGCTTCT 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8860 CTACCAGGAGCGCCCTGCCCTGAGAGAGTCGCCCAGAGAGCCAGAATGTCACTCGGCTTCT 8919

Qy   1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8920 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 8979

Qy   1533 CTGGCAGAGAAAAATAGATGAGACCCTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8980 CTGGCAGAGAAAAATAGATGAGACCCTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 9039

Qy   1593 TGAGCTGGACCTCAAGCTCGCCGCCAAGCTGAGGTGATCAAGGAGTCGGCAGCCGTGGG 1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9040 TGAGCTGGACCTCAAGCTCGCCGCCAAGCTGAGGTGATCAAGGAGTCGGCAGCCGTGGG 9099

Qy   1653 CGATCTCTCATTGACTCTTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
          |||||| |||||||||||||| |||||||||||||||||||||||||||||||||||
Db   9100 CGATCTCCTCATTGACTCTTCCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 9159

Qy   1713 AATTGCGCCTGAAAAGAAGAACGTGAGCCAGTCAATGACTTGCTCGCCAGCTTACCAC 1772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9160 AATTGCGCCTGAAAAGAAGAACGTGAGCCAGTCAATGACTTGCTCGCCAGCTTACCAC 9219

Qy   1773 TTTGGGCATTCAGCTCTCACCGTATAACCTGACACTCTGGAAGACCTGAACACCAGATG 1832
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9220 TTTGGGCATTCAGCTCTCACCGTATAACCTGACACTCTGGAAGACCTGAACACCAGATG 9279

Qy   1833 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTCGCATGAAGCCCACGGGA 1892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9280 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTCGCATGAAGCCCACGGGA 9339

Qy   1893 CTTTGGTCCAGCATCTCAGCACTTCTTCTTCCAGGTCTCGTCCAGGGTCCTGGGAGAGAGC 1952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9340 CTTTGGTCCAGCATCTCAGCACTTCTTCTTCCAGGTCTCGTCCAGGGTCCTGGGAGAGAGC 9399

Qy   1953 CATCTCGCCAAACAAAAGTGCCCTACTATATCAACCAGAGACTCAAACA 2001
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   9400 CATCTCGCCAAACAAAAGTGCCCTACTATATCAACCAGAGACTCAAACA 9448


RESULT 15
ABKB2002
ID   ABKB2002 standard; DNA; 11443 BP.
XX
AC   ABKB2002;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaR9-R16.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200229056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US011126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 17; 145pp; English.
```

XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
XX   invention
SQ

Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 U; 0 Other;

Query Match   59.1%;   Score 1182.6;   DB 6;   Length 11443;
Best Local Similarity   99.7%;   Pred. No. 0;
Matches 1185;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;

```
Qy    813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTCTTGCCTG  872
          ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5746 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTCTTGCCTG 5805

Qy    873 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT  932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5806 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 5865

Qy    933 CCTAGAAGACTCCAAGGAGTAAAAGAGCTAAAGAAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA  992
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5866 CCTAGAAGACTCCAAGGAGTAAAAGAGCTAAAGAAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 5925

Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAG 1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5926 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAG 5985

Qy   1053 ATCCCTGGAAGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5986 ATCCCTGGAAGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT 6045

Qy   1113 CAAGTGGAGTGAACTTCGGAAGAGTCTCTCAACATTAGGTTCCCATTTGGAAGCCAGTCC 1172
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6046 CAAGTGGAGTGAACTTCGGAAGAGTCTCTCAACATTAGGTTCCCATTTGGAAGCCAGTCC 6105

Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTCTGGTGTGGTGGCTACAGCTGAA 1232
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   6106 TGACCAGTGGAAGCGTCTGCACCTTCTGGTGTGGTGGCTACAGCTGAA 6165

Qy   1233 AGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 1292
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6166 AGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 6225

Qy   1293 GAACGATGTACATAGGGGCCTTCAAGAGAGGAATTGAAAACTAAAGAACCTGAATCATGAG 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6226 GAACGATGTACATAGGGGCCTTCAAGAGAGGAATTGAAAACTAAAGAACCTGAATCATGAG 6285

Qy   1353 TACTCTTGAGACTGTACGAATATTTCGACAGACCCTTGAAAGACTTCAGGAAGACTAGAGAAACT 1412
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6286 TACTCTTGAGACTGTACGAATATTTCGACAGACCCTTGAAAGACTTCAGGAAGACTAGAGAAACT 6345

Qy   1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCT 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6346 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCT 6405

Qy   1473 ACGAAAGCAGGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 1532
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6406 ACGAAAGCAGGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGA 6465

Qy   1533 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAGGAACTTCAGAGGCCACGGA 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6466 CTGCAGAGAAAAAATAGATGAGACCCTTGAAAGACTTCAGGAACTTCAGAGGCCACGGA 6525

Qy   1593 TGAGCTGGACCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGCAGCCGTGGG 1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6526 TGAGCTGGACCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGCAGCCGTGGG 6585

Qy   1653 CGATCTCCTCATTGACTTCTTCCAAGATCACCTGGAAAGTCAAGGCACTTCGAGGAGA 1712
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6586 CGATCTCCTCATTGACTTCTTCCAAGATCACCTGGAAAGTCAAGGCACTTCGAGGAGA 6645

Qy   1713 AATTGCGGCCTGAAAGAGAACGTGAGCCACGTGCACCTTGCTCGCGCCAGCTTACCAC 1772
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6646 AATTGCGGCCTGAAAGAGAACGTGAGCCACGTGCACCTTGCTCGCGCCAGCTTACCAC 6705

Qy   1773 TTTGGGCATTCAGTCTCCACGGTATAACCTCAGCACTCTGGAAGACCTGACACCAGATG 1832
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6706 TTTGGGCATTCAGTCTCCACGGTATAACCTCAGCACTCTGGAAGACCTGACACCAGATG 6765

Qy   1833 GAAGCTTCTGCGAGGTGGCCGTCGAGGACCGAGTGCATGAAGCCACAGGGA 1892
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6766 GAAGCTTCTGCGAGGTGGCCGTCGAGGACCGAGTGCATGAAGCCCACACAGGGGA 6825

Qy   1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTCGTGTCCAGGGTCCCTGGGAGAGC 1952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6826 CTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTCGTGTCCAGGGTCCCTGGGAGAGC 6885

Qy   1953 CATCTCGCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
          ||||||||       |||||||||||||||||||||||||||||||||
Db   6886 CATCTCGCCAAACAAACAAGTGCCCTACTATATCAACCACGAGACTCAAACA 6934
```

Search completed: April 4, 2004, 18:19:39
Job time : 529.337 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

f nucleic - nucleic search, using sw model

in on: April 4, 2004, 12:04:07 ; Search time 5216.76 Seconds
                                       (without alignments)
                                       16625.161 Million cell updates/sec

itle: US-09-845-416-6_COPY_1000_3000
erfect score: 2001
equence: 1 ggcagttcattgatggagag.........tcaaccacgagagactcaaaca 2001

coring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

earched: 3470272 seqs, 2167516995 residues

otal number of hits satisfying chosen parameters: 6940544

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries

tabase :   GenEmbl:*
             1: gb_ba:*
             2: gb_htg:*
             3: gb_in:*
             4: gb_om:*
             5: gb_ov:*
             6: gb_pat:*
             7: gb_ph:*
             8: gb_pl:*
             9: gb_pr:*
            10: gb_ro:*
            11: gb_sts:*
            12: gb_sy:*
            13: gb_un:*
            14: gb_vi:*
            15: gb_ba:*
            16: em_fun:*
            17: em_hum:*
            18: em_in:*
            19: em_mu:*
            20: em_om:*
            21: em_or:*
            22: em_ov:*
            23: em_pat:*
            24: em_ph:*
            25: em_pl:*
            26: em_ro:*
            27: em_sts:*
            28: em_un:*
            29: em_vi:*
            30: em_htg_hum:*
            31: em_htg_inv:*
            32: em_htg_other:*
            33: em_htg_mus:*
            34: em_htg_pln:*
            35: em_htg_rod:*
            36: em_htg_mam:*
            37: em_htg_vrt:*
            38: em_sy:*
            39: em_htgo_hum:*
            40: em_htgo_mus:*
            41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	1283	64.1	5462	6	AX538621	AX538621 Sequence
2	1182.6	59.1	5952	6	AR304538	AR304538 Sequence
3	1182.6	59.1	5952	6	AX114289	AX114289 Sequence
4	1182.6	59.1	8689	6	AX538622	AX538622 Sequence
5	1182.6	59.1	11443	6	AX538624	AX538624 Sequence
6	1182.6	59.1	12057	6	AX538627	AX538627 Sequence
7	1182.6	59.1	12446	9	X14298	X14298 Human mRNA
8	1182.6	59.1	13957	6	AX409637	AX409637 Sequence
9	1182.6	59.1	13957	6	AX538581	AX538581 Sequence
10	1182.6	59.1	13957	9	HSDMDR	M18533 Homo sapien
11	1182.6	59.1	13977	6	AR220819	AR220819 Sequence
12	1180	59.0	5339	6	AX538620	AX538620 Sequence
13	1030.6	51.5	13887	4	AF070485	AF070485 Canis fam
14	1004	50.2	13815	6	AX306153	AX306153 Sequence
15	1004	50.2	13815	6	AX538582	AX538582 Sequence
16	1004	50.2	13815	10	MUSDYSA	M68859 Mouse dystr
17	1004	50.2	19307	6	AR093392	AR093392 Sequence
18	1004	50.2	19307	6	AR142592	AR142592 Sequence
19	765	38.3	13575	5	GGDYS	X13369 Chicken mRN
20	727	36.3	5417	6	AX538619	AX538619 Sequence
21	678.4	33.9	4402	6	E30220	E30220 Shortened d
22	618.8	30.7	4075	6	E30221	E30221 Shortened d
23	614.2	30.7	4402	6	E30219	E30219 Shortened d
24	547.4	27.4	3275	10	MUSDYS	M18025 Mouse dystr
25	459.4	23.0	630	9	HSDMDF1	X06179 Human fetal
26	422.6	21.1	3521	5	AF339031	AF339031 Danio rer
27	414.6	20.7	1481	9	BC009242	BC009242 Homo sapi
28	408.2	20.4	11096	6	AX538584	AX538584 Sequence
29	408.2	20.4	11096	10	MMY12229	Y12229 M.musculus
30	408	20.4	10705	10	RNAJ2967	AJ002967 Rattus no
31	397	19.8	6045	6	A63605	A63605 Sequence 7
32	397	19.8	6045	6	AR281528	AR281528 Sequence
33	397	19.8	6059	6	AX107972	AX107972 Sequence
34	397	19.8	10302	6	AX538583	AX538583 Sequence
35	397	19.8	10302	9	HSMOPS	X69086 H.sapiens m
36	397	19.8	10320	6	A63607	A63607 Sequence 9
37	397	19.8	10320	6	AR281529	AR281529 Sequence
38	391.8	19.6	2654	5	FSCDYSTRO	FSCDYSTRO Torpedo cal
39	387	19.3	387	6	AX538612	AX538612 Sequence
40	378.6	18.9	3161	10	MMGUTRPH	MMGUTRPH M.musculus
41	348	17.4	348	6	AX538611	AX538611 Sequence
42	330	16.5	333	6	AX538589	AX538589 Sequence
43	327	16.3	327	6	AX538588	AX538588 Sequence
44	324	16.2	324	6	AX538613	AX538613 Sequence
45	322.6	16.1	3499	9	HSU43519	U43519 Human dystr

ALIGNMENTS

RESULT 1
AX538621
LOCUS   AX538621   5462 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 41 from Patent WO0229056.
ACCESSION   AX538621
VERSION   AX538621.1   GI:25271168
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE   Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 41 11-APR-2002;
             THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

Pred. No. is the number of results predicted by chance to have a

FEATURES Location/Qualifiers
source 1..5462
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:3630"
/note="Synthetic"

ORIGIN

Query Match 64.1%; Score 1283; DB 6; Length 5462;
Best Local Similarity 99.3%; Pred. No. 6.9e-311;
Matches 1302; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy 691 GCTCCTGGACTGACCACTATTGGAGCCCTTCCTACTCAGACTGTTACTCTGGTGACACAA 750
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1547 GCTCCTGGACTGACCACTATTGGAGCCCTTCCTACTCAGACTGTTACTCTGGTGACACAA 1606

Qy 751 CCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTG 810
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1607 CCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTG 1666

Qy 811 GAGGTACCTACTACTACATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTGCC 870
|||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 1667 GAG--------CATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTGCC 1717

Qy 871 TGGCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGG 930
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1718 TGGCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGG 1777

Qy 931 CTCCTAGAAGACTTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACTTCCAAGGT 990
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1778 CTCCTAGAAGACTTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACTTCCAAGGT 1837

Qy 991 GAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCAAAAAATCCTG 1050
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1838 GAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCAAAAAATCCTG 1897

Qy 1051 AGATCCCTGAAAGTTCCGATGATGCAGTCCGTTTGGATAACATGAAC 1110
||||||||||||||||||||||||||||||||||||||||||||||
Db 1898 AGATCCCTGGAAAGTTCCGATGATGCAGTCCGTTTGGATAACATGAAC 1957

Qy 1111 TTCAAGTGGAGTGAGACTTCGGAAAAAGTCTCAACATTAGTCCCATTGGAAGCCAGT 1170
||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1958 TTCAAGTGGAGTGAGACTTCGGAAAAAGTCTCAACATTAGTCCCATTGGAAGCCAGT 2017

Qy 1171 TCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGAGGAACTTCGGTGTGGCTACAGCTG 1230
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1958 TCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGAGGAACTTCGGTGTGGCTACAGCTG 2077

Qy 1231 AAAGATGATGAATTAAGCCCGGCAGGCACCTATTGGAGCGCGACTTCCAGCAGTTCAGAAG 1290
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2078 AAAGATGATGAATTAAGCCCGGCAGGCACCTATTGGAGCGCGACTTCCAGCAGTTCAGAAG 2137

Qy 1291 CAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATG 1350
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2138 CAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATG 2197

Qy 1351 AGTACTCCTTGACAATAGATGAGACCCTTGACGAAGACTGTAGCCGATGATGTCTAAGGAAAGGCT 1410
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2198 AGTACTCTTGACAATAGATGAGACCCTTGACGAAGACTGTAGCCGATGATGTCTAAGGAAAGGCT 2257

Qy 1411 CTCTACCAGGAGCTGCCTCCTGAGAGAGAGCCCAGAGAGCCCAGAATGTCACTCGGCTT 1470
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2258 CTCTACCAGGAGCTGCCTCCTGAGAGAGAGCCCAGAGAGCCCAGAATGTCACTCGGCTT 2317

Qy 1471 CTTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCT 1530
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2318 CTTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCT 2377

Qy 1531 GACTGGCAGAGAGAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACG 1590
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2378 GACTGGCAGAGAGAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACG 2437

Qy 1591 GATGAGCTGACCTGCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCCCGTG 1650

Db 2438 GATGAGCTGACCTGCGCCCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTG 2497

Qy 1651 GGGGATCCTCCTCCATTGACTTGAGCCTCTCCAAGATCACCTCGGAAAGTCAAGGCACTTCGAGGA 1710
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2498 GGCGATCCTCCTCCATTGACTTGAGCCTCTCCAAGATCACCTCGGAAAGTCAAGGCACTTCGAGGA 2557

Qy 1711 GAAATTGCGCCTTCTGAAAGAGAAGAACGTGAGCAACGTGAGCCACGTCAGTGTCTGCTTGCTCGCCAGCTTACC 1770
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2558 GAAATTGCGCCTTCTGAAAGAGAAGAACGTGAGCAACGTGAGCCACGTCAGTGTCTGCTTGCTCGCCAGCTTACC 2617

Qy 1771 ACTTTGGGCATTCAGCTTCTCACGGTATAACCTCAGCACTCTGAAGACCTGAACACCAGA 1830
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2618 ACTTTGGGCATTCAGCTTCTCACGGTATAACCTCAGCACTCTGAAGACCTGAACACCAGA 2677

Qy 1831 TGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCGCATGAGAAGCCCACAGG 1890
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2678 TGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCGCATGAGAAGCCCACAGG 2737

Qy 1891 GACTTTGGTCCAGCATCTCCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGGAGAGA 1950
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2738 GACTTTGGTCCAGCATCTCCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGGAGAGA 2797

Qy 1951 GCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2798 GCCATCTCCGCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2848

RESULT 2
AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced
peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1..5952
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 59.1%; Score 1182.6; DB 6; Length 5952;
Best Local Similarity 99.7%; Pred. No. 1.1e-285;
Matches 1185; Conservative 0; Mismatches 0; Indels 4; Gaps 0;

Qy 813 GGTACCTACTACTACATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 872
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2946 GGAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 3005

Qy 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCCAATGTCCTACAGGATGATGTCTAAGGAAAGGCT 932
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3006 GCTTACAGAAGCTGAAACAACTGCCAATGTCCAATGTCCTACAGGATGATGTCTAAGGAAAGGCT 3065

Qy 933 CCTAGAAGACTTCCAAGGAGTAAAAGAGTGATGAAACAATGCAAGACTTCCAAGGTGA 992
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3066 CCTAGAAGACTTCCAAGGAGTAAAAGAGTGATGAAACAATGCAAGACTTCCAAGGTGA 3125

Qy 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCAAAAAAATCCTGAG 1052
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3126 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAACAGCCAAAAAAATCCTGAG 3185

Qy 1053 ATCCCTGGAAAGTTCCGATGATGCAGTCCGTTGTTACAAAGACGTTTGGATAACATGAACTT 1112
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3186 ATCCCTGGAAAGTTCCGATGATGCAGTCCGTTGTTACAAAGACGTTTGGATAACATGAACTT 3245

```
AUTHORS     Paul,X.L. and Xiao,X.
TITLE       Method and vector for producing and transferring trans-spliced
            peptides
JOURNAL     Patent: WO 0129243-A 1 26-APR-2001;
            DALHOUSE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
            SYSTEM OF HIGHER EDUCATION (US)
FEATURES             Location/Qualifiers
     source          1..5952
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    2897..2898
                     /note="S4 junction site"
     misc_feature    3198..3199
                     /note="S2 junction site"
ORIGIN

Query Match     59.1%;  Score 1182.6;  DB 6;  Length 5952;
Best Local Similarity   99.7%;  Pred. No. 1.1e-285;
Matches 1185;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;
```

```
Qy    813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG   872
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2946  GGAAGAAACTTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG  3005

Qy    873  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGCTACCCGTAAGGAAAGGCT   932
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3006  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGCTACCCGTAAGGAAAGGCT  3065

Qy    933  CCTAGAAGACTTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGA   992
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3066  CCTAGAAGACTTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGA  3125

Qy    993  AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAG  1052
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3126  AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAG  3185

Qy   1053  ATCCCTGGAAGGTTCGGATGAGTCCTCGTTACAAGACGTTTGGATAACATGAACTT  1112
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3186  ATCCCTGGAAGGTTCGGATGAGTCCTCGTTACAAGACGTTTGGATAACATGAACTT  3245

Qy   1113  CAAGTGGAGTGAACTTCGGAAAAAGAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  1172
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3246  CAAGTGGAGTGAACTTCGGAAAAAGAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  3305

Qy   1173  TGACCAGTGGAGCGTCTGCACCTTTCTGCAGAGAACTTCTGGTGTGGCTACAGCTGAA  1232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3306  TGACCAGTGGAGCGTCTGCACCTTTCTGCAGAGAACTTCTGGTGTGGCTACAGCTGAA  3365

Qy   1233  AGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA  1292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3366  AGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA  3425

Qy   1293  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3426  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  3485

Qy   1353  TACTCCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGACGACTAGAGAAACT  1412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3486  TACTCCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGACGACTAGAGAAACT  3545

Qy   1413  CTACCAGGAGCCCAGAGAGCTGCCTCCTCGAGGAGAGAGAGCCCAGAAATGTCACTCGGCCTTCT  1472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3546  CTACCAGGAGCCCAGAGAGCTGCCTCCTCGAGGAGAGAGAGCCCAGAAATGTCACTCGGCTTCT  3605

Qy   1473  ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA  1532
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3606  ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA  3665

Qy   1533  CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGA  1592
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3666  CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGA  3725
```

```
y    1113  CAAGTGGAGTGAACTTCGGAAAAAGAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  1172
b    3246  CAAGTGGAGTGAACTTCGGAAAAAGAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  3305

Y    1173  TGACCAGTGGAGCGTCTGCACCTTTCTGCAGAGAACTTCTGGTGTGGCTACAGCTGAA  1232
b    3306  TGACCAGTGGAGCGTCTGCACCTTTCTGCAGAGAACTTCTGGTGTGGCTACAGCTGAA  3365

Y    1233  AGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA  1292
b    3366  AGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA  3425

Y    1293  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
b    3426  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  3485

Y    1353  TACTCCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGACGACTAGAGAAACT  1412
b    3486  TACTCCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGACGACTAGAGAAACT  3545

Y    1413  CTACCAGGAGCGCCCAGAGAGCTGCCTCCTCGAGGAGAGAGAGCTGCCTCCTCGAGGAGAGAGA  1472
b    3546  CTACCAGGAGCGCCCAGAGAGCTGCCTCCTCGAGGAGAGAGAGCTGCCTCCTCGAGGAGAGAGA  3605

Y    1473  ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA  1532
b    3606  ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA  3665

Y    1533  CTGGCAGAGAAAAATAGATGAGACCCTGAAAGACTCTCCAGGAGGCCACCGA  1592
b    3666  CTGGCAGAGAAAAATAGATGAGACCCTGAAAGACTCTCCAGGAGGCCACCGA  3725

Y    1593  TGAGCTGGACCTCAAGCTGCGCGCAAGCTGAGGTGATCAGGATCCTGGCAGCCCGTGGG  1652
b    3726  TGAGCTGGACCTCAAGCTGCGCGCAAGCTGAGGTGATCAGGATCCTGGCAGCCCGTGGG  3785

Y    1653  CGATCCTCCATTGACTCTCTCCAAGATCACCTCGAGAATCAAGGGCACTTCGAGGAGA  1712
c    3786  CGATCCTCCATTGACTCTCTCCAAGATCACCTCGAGAATCAAGGGCACTTCGAGGAGA  3845

Y    1713  AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTACCAC  1772
c    3846  AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTACCAC  3905

Y    1773  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTCGGAAGAACCTCAGACACCAGATG  1832
c    3906  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTCGGAAGAACCTCAGACACCAGATG  3965

Y    1833  GAAGCTTCTGCAGTGGCCGTCAGGGACCGAGCTGCAGCTGCATGAAGCCCACAGGGGA  1892
c    3966  GAAGCTTCTGCAGTGGCCGTCAGGGACCGAGCTGCAGCTGCATGAAGCCCACAGGGGA  4025

Y    1893  CTTTGGTCCAGCATCCAGCACTTTCTTCTTTCCCAGTCGTCGTCCCAGGGTCCCTGGGAGAGAGC  1952
c    4026  CTTTGGTCCAGCATCCAGCACTTTCTTCTTTCCCAGTCGTCGTCCCAGGGTCCCTGGGAGAGAGC  4085

Y    1953  CATCCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA  2001
b    4086  CATCCTGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA  4134
```

```
RESULT 3
X114289

LOCUS       AX114289     5952 bp     DNA     linear     PAT 11-MAY-2001
DEFINITION  Sequence 1 from Patent WO0129243.
ACCESSION   AX114289
VERSION     AX114289.1     GI:14031259
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
```

```
Db   3232 ATCCCTGGAAGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGATAACATGAACTT 3291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 1172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3292 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC 3351
Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGGAAACTTCTTGGTGTGGCTCACAGCTGAA 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3352 TGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGGAAACTTCTTGGTGTGGCTCACAGCTGAA 3411
Qy   1233 AGATGATGAATTAAGCCGCAGGCACCTATTGGAGGCGACTTCAGCAGTCAGAAGCA 1292
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3412 AGATGATGAATTAAGCCGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCA 3471
Qy   1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3472 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 3531
Qy   1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3532 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 3591
Qy   1413 CTACCAGAGCCCAGAGAGTCTGCCTCCTGAGAGAGCCCAGAGTGTCACTCGGCTTCT 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3592 CTACCAGAGCCCAGAGAGTCTGCCTCCTGAGAGAGCCCAGAGTGTCACTCGGCTTCT 3651
Qy   1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3652 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 3711
Qy   1533 CTGGCCAGAGAAAAATGATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGGGCCACGGA 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3712 CTGGCCAGAGAAAAATGATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGGGCCACGGA 3771
Qy   1593 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGACTCCTGGCAGCCCGTGGG 1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3772 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGACTCCTGGCAGCCCGTGGG 3831
Qy   1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCCGAGATCACCTCGAGGCACTTCAGGAGGA 1712
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3832 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCCGAGATCACCTCGAGGCACTTCAGGAGGA 3891
Qy   1713 AATTGCGCGTCTGAAAGAAGACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3892 AATTGCGCGTCTGAAAGAAGACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3951
Qy   1773 TTTGGGCATTCAGCTCTCACCGTATAACCTCGGAAGACTCTGGAAGACCAGATG 1832
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3952 TTTGGGCATTCAGCTCTCACCGTATAACCTCGGAAGACTCTGGAAGACCAGATG 4011
Qy   1833 GAAGCTTCTGCAGGTGGCCGTGCGAGGACCAGCAGTCAGGCAGCTGCATGAGAGCCCACAGGGA 1892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4012 GAAGCTTCTGCAGGTGGCCGTGCGAGGACCAGCAGTCAGGCAGCTGCATGAGAGCCCACAGGGA 4071
Qy   1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCGTTCGTCCAGGGTCCCTGGGGAGAGC 1952
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4072 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGTCGTTCGTCCAGGGTCCCTGGGGAGAGC 4131
Qy   1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAACA 2001
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   4132 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCAGAGACTCAAACA 4180
```

RESULT 5
AX538624
LOCUS   AX538624   11443 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION Sequence 44 from Patent WO0229056.
ACCESSION   AX538624
VERSION   AX538624.1   GI:25271175
KEYWORDS
SOURCE   synthetic construct
ORGANISM   synthetic construct

RESULT 4
AX538622
LOCUS   AX538622   8689 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION   AX538622
VERSION   AX538622.1   GI:25271171
KEYWORDS
SOURCE   synthetic construct
ORGANISM   synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS   Chamberlain,J.S. and Harper,S.Q.
TITLE   Mini-dystrophin nucleic acid and peptide sequences
JOURNAL   Patent: WO 0229056-A 42 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
source   1..8689
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Synthetic"
ORIGIN

Query Match   59.1%; Score 1182.6; DB 6; Length 8689;
Best Local Similarity   99.7%; Pred. No. 1.1e-285;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy    813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 872
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2992 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 3051
Qy    873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTACAGGATGTCTACCCGTAAGGAAAGGCT 932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3052 GCTTACAGAAGCTGAAACAACTGCCAATGTCTACAGGATGTCTACCCGTAAGGAAAGGCT 3111
Qy    933 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3112 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 3171
Qy    993 AATTGAAGCTCCACACGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG 1052
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3172 AATTGAAGCTCCACACGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG 3231
Qy   1053 ATCCCTGGAAGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGATAACATGAACTT 1112
```

Left column (AX538627, Length 11443)

```
EFERENCE     artificial sequences.
             1
AUTHORS      Chamberlain,J.S. and Harper,S.Q.
TITLE        Mini-dystrophin nucleic acid and peptide sequences
JOURNAL      Patent: WO 0229056-A 44 11-APR-2002;
             THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
EATURES             Location/Qualifiers
  source       1..11443
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Synthetic"
RIGIN
```

```
Query Match    59.1%;   Score 1182.6;   DB 6;   Length 11443;
Best Local Similarity    99.7%;   Pred. No. 1.2e-285;
Matches 1185;   Conservative 0;   Mismatches 4;   Indels 0;   Gaps 0;
```

```
y    813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG   872
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   5746  GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG  5805

y    873  GCTTACAGAAGCTGAAACAACTGCCAATGTCTACGAGATGCTACCCGTAAGGAAAGGCT   932
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   5806  GCTTACAGAAGCTGAAACAACTGCCAATGTCTACGAGATGCTACCCGTAAGGAAAGGCT  5865

y    933  CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCCAAGACCTCCAAGGTGA   992
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   5866  CCTAGAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCCAAGACCTCCAAGGTGA  5925

y    993  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAAATCCTGAG  1052
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   5926  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAAATCCTGAG  5985

y   1053  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT  1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   5986  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT  6045

y   1113  CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTC  1172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6046  CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTC  6105

y   1173  TGACCAGTGGAAGGTCTGCGACCTTCTCTGCGAGGAACTTCGTGTGTGGCTACAGCTGAA  1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6106  TGACCAGTGGAAGGTCTGCGACCTTCTCTGCGAGGAACTTCGTGTGTGGCTACAGCTGAA  6165

y   1233  AGATGATGAATTAAGCCGGCAGCACCTATTGGAGGGGACTTTCCAGCAGTTCAGAAGCA  1292
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6166  AGATGATGAATTAAGCCGGCAGCACCTATTGGAGGGGACTTTCCAGCAGTTCAGAAGCA  6225

y   1293  GAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6226  GAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  6285

y   1353  TACTCTTGAGACTGTACGATATTTCTGACAGAGACAGGCCTTTGAAGGACTAGAGAAACT  1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6286  TACTCTTGAGACTGTACGATATTTCTGACAGAGACAGGCCTTTGAAGGACTAGAGAAACT  6345

y   1413  CTACCAGGAGCGCCAGAGAGCTGCCTCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCT  1472
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6346  CTACCAGGAGCGCCAGAGAGCTGCCTCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCT  6405

y   1473  ACGAAAGCAGGCCGAGGAGCTGGAGGTCAATATGAGTGGGGAAAAATTGAACCTGCACTCCGCTGA  1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6406  ACGAAAGCAGGCCGAGGAGCTGGAGGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTCCGCTGA  6465

y   1533  CTGGCGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGAGGCCACGGA  1592
          |||||||||||||||||||||||||||||||||||||||||||||||||||
b   6466  CTGGCGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGAGGCCACGGA  6525

y   1593  TGAGCTGGACCTCAAGCTGCGCGCAAGCTGGAGGTGATCAAGGATCCTTGGCAGCCCGTGGG  1652
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6526  TGAGCTGGACCTCAAGCTGCGCGCAAGCTGGAGGTGATCAAGGATCCTTGGCAGCCCGTGGG  6585
```

Right column (RESULT 6, AX538627, Length 12057)

```
y   1653  CGATTCCTCATTGACTCTTCTCCAAGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGA   1712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6586  CGATTCCTCATTGACTCTTCTCCAAGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGA   6645

y   1713  AATTGGGCCTCTGAAAGAGAGACG'GAGCCACGTCAATGACCTTGCTCGCAGTTACCAC   1772
          ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
b   6646  AATTGCGCCTCTGAAAGAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTACCAC   6705

y   1773  TTTGGGCATTCAGTCTCCACCGTATAACCTCAGCACTCTGGAAGACTTGGAACACCAGATG  1832
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6706  TTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCACTCTGGAAGACTTGGAACACCAGATG  6765

y   1833  GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA  1892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6766  GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA  6825

y   1893  CTTTGGTCCAGCATCCAGCACTTCTTCCACGTCGTGTCCAGGGTCCCTGGGAGAGAGC  1952
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   6826  CTTTGGTCCAGCATCCAGCACTTCTTCCACGTCGTGTCCAGGGTCCCTGGGAGAGAGC  6885

y   1953  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACAGACTCAAACA  2001
          ||||||||||||||||||||||||||||||||||||||||||||||||
b   6886  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACAGACTCAAACA  6934
```

```
RESULT 6
AX538627
LOCUS        AX538627        12057 bp    DNA    linear    PAT 23-NOV-2002
DEFINITION   Sequence 47 from Patent WO0229056.
ACCESSION    AX538627
VERSION      AX538627.1    GI:25271181
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    Chamberlain,J.S. and Harper,S.Q.
  TITLE      Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL    Patent: WO 0229056-A 47 11-APR-2002;
             THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
  source       1..12057
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Synthetic"
ORIGIN
```

```
Query Match    59.1%;   Score 1182.6;   DB 6;   Length 12057;
Best Local Similarity    99.7%;   Pred. No. 1.2e-285;
Matches 1185;   Conservative 0;   Mismatches 4;   Indels 0;   Gaps 0;
```

```
y    813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG   872
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   8260  GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG  8319

y    873  GCTTACAGAAGCTGAAACAACTGCCAATGTCCAATGTCCTACAGGATGATGATGCTAAGAAAGGCT   932
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   8320  GCTTACAGAAGCTGAAACAACTGCCAATGTCCAATGTCCTACAGGATGATGATGCTAAGAAAGGCT  8379

y    933  CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGA   992
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   8380  CCTAGAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGA   8439

y    993  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG  1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   8440  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG  8499

y   1053  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT  1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   8500  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT  8559
```

```
2y  1113 CAAGTGGAGTGAACTTGGAAAAAGTCTCCAACATTAGGTCCCATTGGAAGCCAGTTC 1172
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  8619 CAAGTGGAGTGAACTTGGAAAAAGTCTCCAACATTAGGTCCCATTGGAAGCCAGTTC 8560

2y  1173 TGACCAGTGGAAGCGTCTGCACCTTCTCGGTCGTGTGGGCTACAGCTGAA 1232
         |||||||||||||||||||||||||||||||||||||||||||||||||
2b  8679 TGACCAGTGGAAGCGTCTGCACCTTCTCGGTCGTGTGGGCTACAGCTGAA 8620

2y  1233 AGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGGAAGCA 1292
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  8739 AGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGGAAGCA 8680

2y  1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGACACCTGTAATCATGAA 1352
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  8799 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGACACCTGTAATCATGAA 8740

2y  1353 TACTCTTGAGACTGTACGAATATTTCTGACACAGCAGCACCTTTGGAAGGACTAGAGAAACT 1412
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  8859 TACTCTTGAGACTGTACGAATATTTCTGACACAGCAGCACCTTTGGAAGGACTAGAGAAACT 8800

2y  1413 CTACCAGAGAGCCCAGAGAGCTGCCTCCTGAGGAGGATCCCAGAGAGTCACTCGGCTTCT 1472
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  8919 CTACCAGAGAGCCCAGAGAGCTGCCTCCTGAGGAGGATCCCAGAGAGTCACTCGGCTTCT 8860

2y  1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTCCGCTGA 1532
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  8979 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTCCGCTGA 8920

2y  1533 CTGGCAGAGAAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGACTTCAAGAGGGCCACCGGA 1592
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  9039 CTGGCAGAGAAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGACTTCAAGAGGGCCACCGGA 8980

2y  1593 TGAGCTGACCTCAAGCTGCGGCCAAGCTGAGCGTGATCAAGGGATCCTGCAGCCCGTGGG 1652
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  9099 TGAGCTGACCTCAAGCTGCGGCCAAGCTGAGCGTGATCAAGGGATCCTGCAGCCCGTGGG 9040

2y  1653 CGATCTCTCAATGACTCTCCAAGATCACCTCGAGAAGTCAGGGCACTTCGAGGAGGA 1712
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  9159 CGATCTCTCAATGACTCTCCAAGATCACCTCGAGAAGTCAGGGCACTTCGAGGAGGA 9100

2y  1713 AATTGCGCCTGTGAAAGAGAACGTGAGCGACTGAGCCTCGTGCTCGCCAGCTACCAC 1772
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  9219 AATTGCGCCTGTGAAAGAGAACGTGAGCGACTGAGCCTCGTGCTCGCCAGCTACCAC 9160

2y  1773 TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 1832
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  9279 TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 9220

2y  1833 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGCCAGTCGATGAGAGCCACCACAGGGGA 1892
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  9339 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGCCAGTCGATGAGAGCCACCACAGGGGA 9280

2y  1893 CTTTGGTCCAGCATCTCCAGCACTTCTTCCACGTCTGTCCCAGGGTCCCCGGGAGAGAGC 1952
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2b  9399 CTTTGGTCCAGCATCTCCAGCACTTCTTCCACGTCTGTCCCAGGGTCCCCGGGAGAGAGC 9340

2y  1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCAACCAGAGACTCAAACA 2001
         |||||||||||||||||||||||||||||||||||||||||||||||||||
2b  9448 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCAACCAGAGACTCAAACA 9400
```

AUTHORS   Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
          Davies,K.E.
TITLE   Two human cDNA molecules coding for the Duchenne muscular dystrophy
          (DMD) locus are highly homologous
JOURNAL   Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE   89345106
PUBMED   2668885
REFERENCE 2 (bases 1 to 12446)
AUTHORS   Rosenthal,A.
TITLE   Direct Submission
JOURNAL   Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
          der DDr, Zentralinstitut fuer Molekularbiologie, Robert-Roessle
          str.10, 1115 Berlin Buch, DDR
COMMENT   see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
          M18533 cDNA were located at x14298 pos.496, 1772, 1965, 2449,
          3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES   Location/Qualifiers
     source   1..12446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X chromosomal, Xp21."
                     /tissue_type="muscle"
                     /dev_stage="fetal and adult."
     CDS   99..11156
                     /note="unnamed protein product; dystrophin (AA 1 - 3685)"
                     /codon_start=1
                     /protein_id="CAA32479.1"
                     /db_xref="GI:30846"
                     /db_xref="GOA:P11532"
                     /db_xref="SWISS-PROT:P11532"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEGKGSTRVHALNNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIILHWQVKNVMKNIMAGLQPTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEEHFQLHHQMHYSQ
                     QITVSLAQGYERTSSPKPRFKSVAYTQAAYVTSDPTRSPFPSQHLEAPEDKSFGSSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH
                     QGRVGNILQLGSKLIGTGKLSEDETEVQEMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLELNDWLTKTEERTRKMEEEPLGPDLEDLKRQVQQHKVLQEDLEQEQVRVN
                     SLTHMVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWVLLQDILLKWQRLTE
                     EQCLFSAWLSEKEDAVNKIHTTGFKVDYEVEHQKIIHLLKKQSMGLYSIK
                     QDLLSTLKNSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTQPSLITQTTVM
                     ETVTVTTREQILVKHAQDELPPPPQKKRQITVDSEIRKRLDVDITELHSWITRSEA
                     VLQSPEFAIFRKEGNFSDIKEKVNAIERRKABKFRKLQDASRSGQALVEQMNVEGVNA
                     DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLEQMTTAENWLKIQ
                     PTTPSEPTAIKSQLKICKDEVNRLGSLQPQIERLKIQSIALKICKGQGPMFLDADFVAF
                     TNHFKQVFSDVQAREKELQTIFDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
                     EIMEQRLGELQALQSSLQEQQSGLYYLSTTVKEMSKKAPSEISRKYQSEFEEIEGRWK
                     KLSSQLVEHCQKLEEQMNKLRKIQNHIQTLKKWMAEVDFLKSEWPALGDSEILLKKQL
                     KQCRILVSDIQTIQPSLNSVNEGGQKIKNEAEPFASRLETELKELNTQWDHMCQQVY
                     ARKEALKGGLEKTVSLQKDLSEMHEWMTQAEEEYLERDFEYKTPDEIQKAFEMKRAK
                     BEAQQKEAKVKLLTESVNSVIAQAPPVAQEALKKELETLTTNVQNLLQSIQSAQETENS
                     NQIRILAQTITDGGVMDELINEELETFNSRWRELHEEAVRRQKLLEQSIQSAQETENS
                     LHLIQESLTFIDKQLAAYIADKVDAAQMQEAQKIQSDLTSHEISLEEMKKHNQGKEA
                     AQRVLSQIDVAQKKLQDVSMKFRLFQKPANFEQRLQESKMILDEVKMHLPALETKSVE
                     QEVVQSQLNHCVNLYKSLSEVKSSVEMVIKTGRQIVQKKQTENPKELDERVTALKLHY
                     NELGAKVTERKQQLEKCLKLSRMRKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
                     VAWGKATQREIEKQKVHLKSITEVGEALKTVLGKKETLVEDKLSLLNSNWIAVTSRAE
                     EWLNLLLEYQKHMETFDQNVDHITKWIIQADTILDESEKKKPQQKEDVLKRLKAELND
                     IRPKVDSTRDQAANLMANHGDHCRKLVEPQISELNHRFAAISRHIKTGKASIPLKELE
                     QFNSDIQKLLEPLEAEIQQGVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQQRITDER
                     KSEBIKIKQQLLQTKHNALKDLRSQRRKALEISHQWYQYKRQADDLLKCLDDIEKKL
                     ASLPBPRDERKIKEIDRELQKKKEELNAVRQABGLSEDGAAMAVEPTIQUSKRWRE
                     IESKPAQFRRLNFAQIHTVREETMVMTEDMPLEISYVPSTYLTETHVSQALLEVEQ
                     LLNARDLCAKDFEDLFKQEESLNIKDSLQQSSGRIDIIHSKKTAALQSATPVERVKL
                     QEALSQLDFQWEKVNKMYKDRQGRFDRSVEKWRRFHYDIKIFNQWLTEAEQPLRKTQI
                     PENWEHAKYKWYLKELQDGIGQRQTVRTLNATGEEIIQQSSKTDASILQEKLGSLNL
                     RWQEVCKQLSDRKKRLEEQKNILSEQRDLNEFVLWLEEADNIASIPLEPGKEQQLKE
                     KLEQVKLLVEELPLRQGILKQLNETGGPVLVSAPISPEEQDKLENKLKQTNLQWIKVS
                     RALPEKQGEIEAQIKDLGQLEKKLEDLEEQLNHLLLWLSPIRNQLEIYNQPNQEGPFD
                     VKETBIAVQAKQPDVEEILSKGQHLYKEKPATQPVKRKLEDLSSEWKAVNRLLQELRA"

RESULT 7
LOCUS   HSDMDR   12446 bp   mRNA   linear   PRI 12-SEP-1993
DEFINITION Human mRNA for dystrophin.
ACCESSION   X14298
VERSION   X14298.1   GI:30845
KEYWORDS   Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 12446)

```
        KQPDLAPGLTTIGASPTQTVLVTQPVTKETAISKLEMPSSLMEVPALADFNRAWT
        ELTDWLSLLDQVIKSQRVMVGDLEDINEMIKQATMQDLEQRRPQLELITAAQNLK
        NKTSNQEARTIITDRIERIQNQWDEYVQEHIQNRRQQLNEMLKDSTQWLBEAKEAEQVL
        GQARAKLESWKBGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSAD
        DTRKVHMITENINASWRSIHKRVSBREAALEETHRLLQQFPLDLEKFLAWLTRAETTA
        NVLIQDATRKERILLEDSKGVKELMKQWQDLQGEIEAHTDVHNLDENSQKILRSLEGSD
        DAVLLQRRLDNNMFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELLWLQLQLKDDEL
        SROAPIGGDFPAVOKQNDVHRAFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLYQ
        EPRELPEERAQNVTRLLRLRKQAEEVNTEWEKLNLHSADWQRKIDETLERIQELQEATD
        ELDLKLRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEIAPLKENVHVNDLARQL
        TTLGIQLSPYNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASOHFLSTSVQGP
        WERAISPNKVPYYINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRLQKA
        LCLDLLSLSAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHNNLVNVPLCVD
        MCLNWLLNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRL
        GLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIEAALFLDMRLEPQSM
        VWLPVLHRVAAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDICQSCFFSGRVAKGH
        KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLEGDNMET
        PVTLINFWPVDSAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSYLNDSISPNESID
        DEHLLIQHYCQSLNQDSFLSQPRSPAQILISESERGELERILADLEBENRNLQAEY
        DRLKQQHEHKGLSPLPSPPEMMPTSPQSPRDAELIABAKLLRQHKGRLEARMQILEDH
        NKQLESQLHRLRQLLEQPQAEAKVNGTTVSSPSTSLRQSDSSQPMLLRVVSGQTSDSM
        GEEDLLSPQDTISTGLBEVMEQLNNSFPSSRGRNTPGKPMREDTM"
```

RIGIN

Query Match   59.1%;   Score 1182.6;   DB 9;   Length 12446;
Best Local Similarity   99.7%;   Pred. No. 1.2e-285;
Matches 1185;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;

```
y    813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG  872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8150  GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 8209

y    873  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT  932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8210  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 8269

y    933  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGA  992
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8270  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGA 8329

y    993  AATTGAAGCTCACACAGATGTTATCACAACCTGAGAAAACATCAGCCAAAAAATCCTGAG 1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8330  AATTGAAGCTCACACAGATGTTATCACAACCTGAGAAAACAGCCAAAAAATCCTGAG 8389

y   1053  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTT 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8390  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTT 8449

y   1113  CAAGTGGAGTGAACTTGGAAAAGTCTCTCAACATTAGGTCCATTGGAAGCCAGTTC 1172
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8450  CAAGTGGAGTGAACTTGGAAAAGTCTCTCAACATTAGGTCCATTGGAAGCCAGTTC 8509

y   1173  TGACCAGTGAAGCTCTGCACCTTTCTCTGCGAACTTCGGTGTGTGGCTACAGCTGAA 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8510  TGACCAGTGAAGCTCTGCACCTTTCTCTGCGAACTTCGGTGTGTGGCTACAGCTGAA 8569

y   1233  AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCACCTTTCCAGCAGTTCAGAAGCA 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8570  AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCACCTTTCCAGCAGTTCAGAAGCA 8629

y   1293  GAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGAACCTGTAATCATGAG 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8630  GAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGAACCTGTAATCATGAG 8689

y   1353  TACTCTTGAGACTGTACGAATATTCTGACAGACAGGCTTTGGAAGGACTAGAGAAACT 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8690  TACTCTTGAGACTGTACGAATATTCTGACAGACAGGCTTTGGAAGGACTAGAGAAACT 8749

y   1413  CTACCAGGAGCCCAGAAGCTGCCTCCCGAGGAGGAGCCCAATGTCACTCGGCTTCT 1472
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
o   8750  CTACCAGGAGCCCAGAAGCTGCCTCCCGAGGAGGAGCCCAATGTCACTCGGCTTCT 8809

y   1473  ACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGAGAAAATTGAACCTGCACTCCGCTGA  992
```

RESULT 8

AX409637
LOCUS   AX409637   13957 bp   DNA   linear   PAT 14-JUN-2002
DEFINITION   Sequence 2284 from Patent WO0229103.
ACCESSION   AX409637
VERSION   AX409637.1   GI:21442342
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE   Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2284 11-APR-2002;
             GENE LOGIC INC (US)
FEATURES   Location/Qualifiers
     source   1..13957
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. M18533"
ORIGIN

Query Match   59.1%;   Score 1182.6;   DB 6;   Length 13957;
Best Local Similarity   99.7%;   Pred. No. 1.2e-285;
Matches 1185;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;

```
Qy    813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG  872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8260  GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 8319

Qy    873  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT  932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8320  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 8379

Qy    933  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGA  992
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8810  ACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGAGAAAATTGAACCTGCACTCCGCTGA 8869

Qy   1533  CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACCGA 1592
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8870  CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACCGA 8929

Qy   1593  TGAGCTGGACCTCAAGCTGCGCGCAAGCTGAGGTGATCAAGCTGGAGCCCGTGGG 1652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8930  TGAGCTGGACCTCAAGCTGCGCGCAAGCTGAGGTGATCAAGCTGGAGCCCGTGGG 8989

Qy   1653  CGATCTCCTCATTGACTCCTCCAAGATCACCACCTGAGAAGTCAAGGACTTGAGGAGA 1712
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8990  CGATCTCCTCATTGACTCCTCCAAGATCACCACCTGAGAAAGTCAAGGCACTTCGAGGAGA 9049

Qy   1713  AATTGCGCCTCTGAAAGAAGAACGTGAGCCACGTCAATGACCTTGCTGCCCAGCTTACCAC 1772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9050  AATTGCGCCTCTGAAAGAAGAACGTGAGCCACGTCAATGACCTTGCTGCCCAGCTTACCAC 9109

Qy   1773  TTTGGGCATTCAGCTTCACCGTATAACCTGCAGTCAGCACACTCAGCACTCTGGAAGACTCAGCATG 1832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9110  TTTGGGCATTCAGCTTCACCGTATAACCTGCAGTCAGCACTCAGCACTCTGGAAGACTCAGCATG 9169

Qy   1833  GAAGTTCTGCAGGTGGCGTGCAGGACCAGGAGTCAGGCGAGTCGCATGAGAGCCCACAGGGA 1892
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9170  GAAGTTCTGCAGGTGGCGTGCAGGACCAGGAGTCAGGCGAGTCGCATGAGAAGCCCACAGGGA 9229

Qy   1893  CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCTGGGAGGAGAGC 1952
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9230  CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCTGGGAGGAGAGC 9289

Qy   1953  CATCTCGCCAAACAAAGTGCCCACTACTATATCAACCACGAGACTCAAACA 2001
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9290  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 9338
```

```
LOCUS       AX538581                13957 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO0229056.
ACCESSION   AX538581
VERSION     AX538581.1  GI:25271086
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 1 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Query Match   59.1%; Score 1182.6; DB 6; Length 13957;
Best Local Similarity   99.7%; Pred. No. 1.2e-285;
Matches 1185; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    813 GGTACCTACTCATCATGAGATTAATACTGCAACAGTCCCCCTGACCAAGTTTCTTGCCTG  872
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8260 GGAAGAAACTCATAGATTATCACAGATTTATCACAACAGTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 8319

Qy    873 GCTTACAGAAGCTGAAACAACTGCCAATGTCCAATGTCCTACAGATGTCTACACCCGTAAGGAAAGGCT  932
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8320 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGTCTACACCCGTAAGGAAAGGCT 8379

Qy    933 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA  992
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8380 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439

Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG 1052
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG 8499

Qy   1053 ATCCCTGGAAGTTCCGATGATGCAGTCCGTGTTACAAAGAGTTTGATAACATGAACTT 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8500 ATCCCTGGAAGTTCCGATGATGCAGTCCGTGTTACAAAGAGTTTGATAACATGAACTT 8559

Qy   1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTTGGAAGCCAGTTC 1172
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   8560 CAAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTTGGAAGCCAGTTC 8619

Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTCTCGCGGAACTTCGTGTGTGGCAGCTGACAGCTGAA 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8620 TGACCAGTGGAAGCGTCTGCACCTTCTCGGTGTGGCAGCTGACAGCTGAA 8679

Qy   1233 AGATGATGAATTAAGCCGGACACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8680 AGATGATGAATTAAGCCGGACACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 8739

Qy   1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTAATCATGAG 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8740 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTAATCATGAG 8799

Qy   1353 TACTCTTGAGACTGTACGAATATTTCTGAAGGACTAGAGAAACT 1412
          ||||||||||||||||||||||||||||||||||||||||||||
Db   8800 TACTCTTGAGACTGTACGAATATTTCTGAAGGACTAGAGAAACT 8859

Qy   1413 CTACCAGAGCCCAGAGAGCGAGAGGAGAGGAGCCCAGAATGTCACTCGGCTTCT 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8860 CTACCAGAGCCCAGAGAGCGAGAGGAGAGGAGCCCAGAATGTCACTCGGCTTCT 8919

Qy   1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTCCGCTGA 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8920 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTCCGCTGA 8979
```

```
Db   8380 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439
Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG 1052
Db   8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG 8499
Qy   1053 ATCCCTGGAAGTTCCGATGATGCAGTCCGTGTTACAAAGAGTTTGATAACATGAACTT 1112
Db   8500 ATCCCTGGAAGTTCCGATGATGCAGTCCGTGTTACAAAGAGTTTGATAACATGAACTT 8559
Qy   1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTTGGAAGCCAGTTC 1172
Db   8560 CAAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTTGGAAGCCAGTTC 8619
Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTCTCGGTGTGGCAGCTGACAGCTGAA 1232
Db   8620 TGACCAGTGGAAGCGTCTGCACCTTCTCGGTGTGGCAGCTGACAGCTGAA 8679
Qy   1233 AGATGATGAATTAAGCCGGACACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 1292
Db   8680 AGATGATGAATTAAGCCGGACACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 8739
Qy   1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTAATCATGAG 1352
Db   8740 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTAATCATGAG 8799
Qy   1353 TACTCTTGAGACTGTACGAATATTTCTGAAGGACTAGAGAAACT 1412
Db   8800 TACTCTTGAGACTGTACGAATATTTCTGAAGGACTAGAGAAACT 8859
Qy   1413 CTACCAGAGCCCAGAGAGCGAGAGGAGAGGAGCCCAGAATGTCACTCGGCTTCT 1472
Db   8860 CTACCAGAGCCCAGAGAGCGAGAGGAGAGGAGCCCAGAATGTCACTCGGCTTCT 8919
Qy   1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTCCGCTGA 1532
Db   8920 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTCCGCTGA 8979

Qy   1533 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGGCCACGGA 1592
Db   8980 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGGCCACGGA 9039
Qy   1593 TGAGCTGACCTCAAGCTGCTCGGCCAAGCTGAGTGATCAAGGATCAAGGCACTCGAGGGA 1652
Db   9040 TGAGCTGACCTCAAGCTGCTCGGCCAAGCTGAGTGATCAAGGATCAAGGCACTCGAGGGA 9099
Qy   1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
Db   9100 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 9159
Qy   1713 AATTTGCGCCTCTGAAAGAACGTGAGCCACTGACCTTGCTCGCCAGCTTACCAC 1772
Db   9160 AATTTGCGCCTCTGAAAGAACGTGAGCCACTGACCTTGCTCGCCAGCTTACCAC 9219
Qy   1773 TTTGGCATTCAGCTCAGCACTTTCTTTTCCACGTATAACCTCGGAAGACTCGAACACCAGATG 1832
Db   9220 TTTGGCATTCAGCTCAGCACTTTCTTTTCCACGTATAACCTCGGAAGACTCGAACACCAGATG 9279
Qy   1833 GAAGCTTCTCGCAGGTGGCCGTCGAGCCGTCGGCAGCTGCATGAGCCCACAGGGA 1892
Db   9280 GAAGCTTCTCGCAGGTGGCCGTCGAGCCGTCGGCAGCTGCATGAGCCCACAGGGA 9339
Qy   1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
Db   9340 CTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 9399
Qy   1953 CATCCTCGCCCAAACAAAGTGCCCTACTATCATCAACCACGAGACTCAAACA 2001
Db   9400 CATCCTCGCCCAAACAAAGTGCCCTACTATCATCAACCACGAGACTCAAACA 9448
```

RESULT 9

```
ly   1533 CTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAGGAGGCCACGGA 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
lb   8980 CTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAGGAGGCCACGGA 9039

ly   1593 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG 1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
lb   9040 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGG 9099

ly   1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
lb   9100 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 9159

ly   1713 AATTGCGCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
lb   9160 AATTGCGCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 9219

ly   1773 TTTGGGCATTCAGCTTCTCCACGTATAACTTCAGCACTCTGGAAGACCTGAACACCAGATG 1832
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
lb   9220 TTTGGGCATTCAGCTTCTCCACGTATAACTTCAGCACTCTGGAAGACCTGAACACCAGATG 9279

ly   1833 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGGAGTCAGGCGAGTCGCATGAAGAGCCACAGGGA 1892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
lb   9280 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGGAGTCAGGCGAGTCGCATGAAGAGCCACAGGGA 9339

ly   1893 CTTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGC 1952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
lb   9340 CTTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAGAGC 9399

ly   1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
          ||||||||||||||||||||||||||||||||||||||||||||||||
lb   9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 9448
```

```
RESULT 10
LUMDYS
LOCUS       HUMDYS          13957 bp    mRNA    linear      PRI 25-MAY-2000
DEFINITION  Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION   M18533 M17154 M18026 M20250
VERSION     M18533.1  GI:181856
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1699)
  AUTHORS   Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and
            Kunkel,L.M.
  TITLE     Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
            preliminary genomic organization of the DMD gene in normal and
            affected individuals
  JOURNAL   Cell 50 (3), 509-517 (1987)
  MEDLINE   87273512
  PUBMED    3607877
REFERENCE   2  (bases 1678 to 3830)
  AUTHORS   Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
  TITLE     Conservation of the Duchenne muscular dystrophy gene in mice and
            humans
  JOURNAL   Science 238 (4825), 347-350 (1987)
  MEDLINE   88018015
  PUBMED    3659917
REFERENCE   3  (bases 1 to 13957)
  AUTHORS   Koenig,M., Monaco,A.P. and Kunkel,L.M.
  TITLE     The complete sequence of dystrophin predicts a rod-shaped
            cytoskeletal protein
  JOURNAL   Cell 53 (2), 219-226 (1988)
  MEDLINE   88194521
  PUBMED    3282674
COMMENT     On May 25, 2000 this sequence version replaced gi:340693.
            Draft entry and computer-readable sequence kindly provided by
            M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
            determined by the size of the deleted DNA segment. Deletions found
            in different patients were from positions 302-2200, 473-1168,
```

```
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xp21.3-p21.1"
                     /tissue_type="muscle"
                     /dev_stage="fetus"
                     /note="G00-119-850"
     gene            1..13957
                     /gene="DMD"
     CDS             209..11266
                     /gene="DMD"
                     /codon_start=1
                     /product="dystrophin"
                     /protein_id="AAA53189.1"
                     /db_xref="GI:181857"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIILHWQVQNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVVINFT
                     TSWSDGLAINALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEEHFQLHQMHYSQ
                     QITVSLAQGYERTSSPKPRFKSYAYTQAAYVTSDPTRSPFPSQHLEAPEDKSFGSSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTIQAQGEISNDVEVVKDQFHTHEGYMDLTAH
                     QGRVGNIIQLQGSKLJGTGKLSEDEETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLKELNDWLTKTEERTRKMEEEPLGPDLEDLKRQVQHKVLQEDLEQGQVRVN
                     SLTHMVVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRMVLLQDILLKWQRLTE
                     EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLQKLAVLKADLEKKKQSMGKLYSLK
                     QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTTQPSLTQTTVM
                     ETVTTVTTREQILVKHAQEELPPPPQKKRQITVDSEIRKRLDVDITELHSWITRSEA
                     VLQSPEFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSAQALVEQMVNEGVNA
                     DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLEQMTTAENWLKIQ
                     PTPSEPTAIKSQLKICKDEVNRLSGLQPQIERLKIOSIALKEKGQQPMFLDADFVAF
                     TNHFKQVFSDVQAREKELQTIFDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
                     EIMEQRLGELQALQSSLQEQQSGLYLSTTVKEMSKKAPSEISRKYQSEFEEIGRWK
                     KLSSQLVEHCQKLEEQMNKLRKIQNHIQTLKKWMAEVDVFLKEEWPALGDSEILKKQL
                     KQCRLLVSDIQTIQPSLNSVNEGGQKIKNEAEPPASRLETELKELNTQWDHMCCQVY
                     ARKEALKGGLEKTVSLQKDLSEMHEWMTQABEEYLERDFEYKTPDELQKAVEEMKRAK
                     EEAQQKEAKVKLLTESVNSVIACAPPVAQEALKKELETLTTNYQWLCTRLNGKCKTLE
                     EVWACWHELLSYLEKANKWLNEVFEKLKTTENIPGGAEISEVLDSLENLMRHSEDNP
                     NQIRILAQTIDGVMDEIINEBLETFNSRWRELHEEAVRRQKLLEQSIQSAQETEKS
                     LHLIQESLTFIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLEMKKHNQGKEA
                     AQRVLSQIDVAQRKLQDVSMKFRLFQKPANFELRLQESKMILDEKVKMVHLPALETKSVE
                     QEVVSQSLNHCVNLYKSLSEVKSEVEMVIKTGRQIVQKKQTENPKELDERVTALKLHY
                     NELGAKVTERKQQLEKCLKLSRKMRKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
                     VAWGKATQKEIEKQKVHLKSITEVGEALKTVLGKKETIVEDLKLSNLNSNWIAVTSRAE
                     RWDEVCKQLSDRKRKLBEQKNILSEFQRDLNEFVLWLEEADNIASIPLEPGKEQQLKE
                     KLEQVKLLVEELPLRQGIIKQLNETGGPVLVSAPISPEEQDKLENKIKQTNLQWIKVS
                     RALPEKQGIEAQLDLGLEKKLEDLEEQLNHLLLWLSPIRNQLEIYNQPNQEGPFD
                     VQETEIAVQAKQPDVEEILSKQQHLYKEKPATQPVKRLKLEDLSSEWKAVNRLLQELRA
                     KQPDLAPGLTTIGASPTQTVTLVTQPVVTKETAISKLEMPSSLMLEVPALADFNRAWT
                     ELTDWLSLLDQVIKSQRVMVGDLEDINEMIIKQKATMQDLEQRRPQLEELITAAQNIK
                     NKTSNQEARTIITDRIERIQNOWDEVCEHLQKDLRQWQTNVDVANDLALALKLRDYSAD
                     GQARAKLSWKEGPYTVDAIQKKITETKQLAKDLRQWQTNVVANDLALKLRGEIAPLKENVSHVNLARQL
                     DTRKVHMITENINASWRSIHKRVSEREAALEETHRLLQQFPLDLEKFLAWLTEAETTA
                     NVLQDATRKERLLEDSKGVKELMKQWQDLQGEIEAHTDVYHNLDENSQKILRSLEGSD
                     DAVLLQRLDNMNFKWSELRKKSLNIRSHLEASSDOWKRLHLSSLQLLVWLQLKDDEL
                     SRQAPIGGDFPAVQKQNDVHRAFKRELKTKEPVIMSLTEVRIFLITEQPLEGLEKLYQ
                     EPRELPPERAQNVTRLLRKQAPENTWEKLNLHSADWQRKIDETLERLQEBATD
                     ELDKLRQAEVIKGSWQPVGDLLIDSLQDLKEVKALRGEIAPLKENVSHVNDLARQL
                     TTLGIQLSPYNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVCQP
                     WERAISPNKVPYYINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA
```

```
              LCLDLSLSAACDALDQHNLKQNDQPMDIIQIINCLTTIYDRLEQEHNNLVNVPLCVD
              MCLNWLLNVYDFGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKVVASSTGFCDQRRL
              GLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIEAALFLDMRLEPQSM
              VNIPVLHRVAAETAKHQAKKCNICKECPIIGFRYRSLKHFNYDICQSCFFSGRVAKGH
              KNHYPMVEYCTPTTSGBDVRDFAKVLNKFRTKYPAKGHPRMGYLPVQTVLEGONMET
              PVTLINFWPVDSAPASSPQLSHDDTHSRIFEHYASRLAEMENSNGSYLNDSISPNESID
              DEHLLIQHYCQSLNQDSPLSQPRSPAQIILSLESSEBERGELERIIADLLEBENRNIQAEY
              DRIKQQHEHKGISPLPFPEMMPTSPSPROAELIAEAKLLRQHGRLRARMQLLEDH
              NKQLESQLHRLRQLLEQPQAEAKVNGTTVSSPSTSLQRSDSSSQPMLLRVVGSQTSDSM
              GBEDLLSPPQDTSTGLEEVMSQLNNSFPSSRGRNTPGKPMREDTM"
ORIGIN

Query Match    59.1%;  Score 1182.6;  DB 9;  Length 13957;
Best Local Similarity 99.7%;   Pred. No. 1.2e-285;
Matches 1185;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy    813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG  872
          ||||  ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 8319

Qy    873 GCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGATGTACCCGTAAGGAAAGGCT  932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8320 GCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACGGATGTCCTACAGATGTACCCGTAAGGAAAGGCT 8379

Qy    933 CCTAGAAGACTCCAAGGGAGTAAAAGACGTGATGAAACAATGCAAGACCTCCAAGGTGA  992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8380 CCTAGAAGACTCCAAGGGAGTAAAAGAGCTAAGTGCTGATGAGTGTAAAAGAGCTAAGTGCTGATGAGTGTAAAAGACGTGATGAAACAATGCAAGACCTCCAAGGTGA 8439

Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG 1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG 8499

Qy   1053 ATCCCTGGAAGGTTCGGATGATGCAGTCCGATGATGCAGTTGGATAACATGAACTT 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8500 ATCCCTGGAAGGTTCGGATGATGCAGTCCGATGATGCAGTTGGATAACATGAACTT 8559

Qy   1113 CAAGTGGGAGTTGAACTTCGGAAAAAGTCCCATTGGAAGCCAGTTC 1172
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   8560 CAAGTGGGAGTTGAACTTCGGAAAAAGTCCCATTGGAAGCCAGTTC 8619

Qy   1173 TGACCAGTGAAGCGTCTGCACCTTCTGGTGTGGCTACAGCTGAA 1232
          ||||||||||||||||||||||||||||||||||||||||||||
Db   8620 TGACCAGTGAAGCGTCTGCACCTTCTGGTGTGGCTACAGCTGAA 8679

Qy   1233 AGATGATGAATTAAGCCGGCAGCCGCCGGACCTATTGGAGGGCGACTTTCCAGCAGTTCAGAAGCA 1292
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8680 AGATGATGAATTAAGCCGGCAGCCGCCGGACCTATTGGAGGGCGACTTTCCAGCAGTTCAGAAGCA 8739

Qy   1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTGTTAAAGAACCTGTAATCATGAG 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8740 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAATCATGAG 8799

Qy   1353 TACTCTTGAGACTGTACGAATGTTGGAAGGACTAGAGAAACT 1412
          ||||||||||||||||||||||||||||||||||||||||||
Db   8800 TACTCTTGAGACTGTACGAATGTTGGAAGGACTAGAGAAACT 8859

Qy   1413 CTACCAGGAGCCCAAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTCT 1472
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8860 CTACCAGGAGCCCAAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTCT 8919

Qy   1473 ACGAAAGCAGCCGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCAC 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8920 ACGAAAGCAGCCGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTCCGCTGA 8979

Qy   1533 CTGGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGA 1592
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8980 CTGGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGA 9039

Qy   1593 TGAGCTGGACCTCAAGCTGGACTTGATCAAGGAGATCCTGGCAGCCCGTGGG 1652
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9040 TGAGCTGGACCTCAAGCTGGACTTGATCAAGGAGATCCTGGCAGCCCGTGGG 9099

Qy   1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTGCAGTCGCGTGGCCACGAGTCAAGCAAGCTGCACTTCGAGGAGA 1712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9100 CGATCTCCTCATTGACTCTCTCCAAGATCACCTGCAGTCGCGTGGCCACGAGTCAAGCACTCGAGGAGA 9159

Qy   1713 AATTGCGCCCTCGAAAAGAACGTGAGCCAGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9160 AATTGCGCCCTCGAAAAGAACGTGAGCCAGTCAATGACCTTGCTCGCCAGCTTACCAC 9219

Qy   1773 TTTGGGCATTCAGCTCACCGTATAACCTCAGCACTCTGGAGACTGAACACCAGAGTG 1832
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9220 TTTGGGCATTCAGCTCACCGTATAACCTCAGCACTCTGGAGACTGAACACCAGCAGATG 9279

Qy   1833 GAAGCTTCTGCAGTGGCCGTGCAGGACCGAGTCAGGCAGTGCATGAAGCCCACAGGGA 1892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9280 GAAGCTTCTGCAGTGGCGGTCCGTCCGAGGACGAGTCAGGCAGTGCATGAAGCCCACAGGGA 9339

Qy   1893 CTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9340 CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC 9399

Qy   1953 CATCTCCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
          |||||| |||||||||||||||||||||||||||||||||||||||||
Db   9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 9448


RESULT 11
AR220819
LOCUS       AR220819              13977 bp    DNA     linear   PAT 26-SEP-2002
DEFINITION  Sequence 60 from patent US 6426186.
ACCESSION   AR220819
VERSION     AR220819.1  GI:23327696
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 13977)
  AUTHORS   Jones,K.A., Volkmuth,W. and Walker,M.G.
  TITLE     Bone remodeling genes
  JOURNAL   Patent: US 6426186-A 60 30-JUL-2002;
FEATURES             Location/Qualifiers
     source          1..13977
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match    59.1%;  Score 1182.6;  DB 6;  Length 13977;
Best Local Similarity 99.7%;   Pred. No. 1.2e-285;
Matches 1185;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy    813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG  872
          ||||  ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 8319

Qy    873 GCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGATGTACCCGTAAGGAAAGGCT  932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8320 GCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACAGATGTACCCGTAAGGAAAGGCT 8379

Qy    933 CCTAGAAGACTCCAAGGGAGTAAAAGACGTGATGAAACAATGCAAGACCTCCAAGGTGA  992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8380 CCTAGAAGACTCCAAGGGAGTAAAAGACGTGATGAAACAATGCAAGACCTCCAAGGTGA 8439

Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG 1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG 8499

Qy   1053 ATCCCTGGAAGGTTCGGATGATGCAGTTGGATAACATGAACTT 1112
          |||||||||||||||||||||||||||||||||||||||||||
Db   8500 ATCCCTGGAAGGTTCGGATGATGCAGTTGGATAACATGAACTT 8559

Qy   1113 GCTTACAGAAGCTGAAAACAACTGCCAATGTCCAATGTCCTACAGATGTACCCGTAAGGAAAGGCT 1172
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   8320 GCTTACAGAAGCTGAAAACAACTGCCAATGTCCAATGTCCTACAGATGTACCCGTAAGGAAAGGCT 8379

Qy    933 CCTAGAAGACTCCAAGGGAGTAAAAGACGTGATGAAACAATGCAAGACCTCCAAGGTGA  992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8380 CCTAGAAGACTCCAAGGGAGTAAAAGACGTGATGAAACAATGCAAGACCTCCAAGGTGA 8439

Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG 1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG 8499

Qy   1053 ATCCCTGGAAGGTTCGGATGATGCAGTTTGGATAACATGAACTT 1112
          ||||||||||||||||||||||||||||||||||||||||||||
Db   8500 ATCCCTGGAAGGTTCGGATGATGCAGTTTGGATAACATGAACTT 8559

Qy   1113 CAAGTGGGAGTTGAACTTCGGAAAAAGTCCCATTGGAAGCCAGTTC 1172
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   8560 CAAGTGGGAGTTGAACTTCGGAAAAAGTCCCATTGGAAGCCAGTTC 8619

Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTCTGCACCTTTCTGCGAGAACTTCGGTGGTGGCTACAGCTGGTGGCTACAGCTGAA 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   8620 TGACCAGTGGAAGCGTCTGCACCTTTCTGCGAGAACTTCGGTGGTGGCTACAGCTGAA 8679

Qy   1293 AATTGAAGCTCACACAGATGTTTATCACAACCTGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG 1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG 8499

Qy   1052 ATCCCTGGAAGGTTCCGATGATGCAGTTTACAGGTTCCGATGATGCAGTTTACAGATAACATGAACTT 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8500 ATCCCTGGAAGGTTCCGATGATGCAGTTTACAGGTTCCGATGATGCAGTTTACAGATAACATGAACTT 8559

Qy   1172 CAAGTGGGAGTTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTC 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8560 CAAGTGGGAGTTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTC 8619

Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTTCTGCGAGAACTTCGGTGGTGGCTACAGCTGAA 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   8620 TGACCAGTGGAAGCGTCTGCACCTTTCTGCGAGAACTTCGGTGGTGGCTACAGCTGAA 8679
```

```
                                         /organism="synthetic construct"
                                         /mol_type="unassigned DNA"
                                         /db_xref="taxon:32630"
                                         /note="Synthetic"

ORIGIN

          Query Match         59.0%;  Score 1180;  DB 6;  Length 5339;
          Best Local Similarity  100.0%;  Pred. No. 5e-285;
          Matches 1180;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    822 TCATAGATTACTGCACAGTCCCCTGACCTGAAAAAGTTTCTGCCTGGCTTACAGA 881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 TCATAGATTACTGCACAGTCCCCTGACCTGAAAAAGTTTCTGCCTGGCTTACAGA 1605

Qy    882 AGCTGAAACAACTGCCAATGTCCTACAGATGCTACCCGTAAGGAAAAGCTCCTAGAAGA 941
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 AGCTGAAACAACTGCCAATGTCCTACAGATGCTACCCGTAAGGAAAAGCTCCTAGAAGA 1665

Qy    942 CTCCAAGGGAGTAAAAGAGAGCTGAGAAACAATGCAAGACCTCCAAGGTGAAATTGAAGC 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 CTCCAAGGGAGTAAAAGAGAGCTGAGAAACAATGCAAGACCTCCAAGGTGAAATTGAAGC 1725

Qy   1002 TCACACAGATGTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1061
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1726 TCACACAGATGTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAGATCCCTGA 1785

Qy   1062 AGGTTCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1786 AGGTTCCGATGATGCAGTCCTGTTACAAACGTTTGGATAACATGAACTTCAAGTGGAG 1845

Qy   1122 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1846 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1905

Qy   1182 GAAGGCTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAAGATGATGA 1241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1906 GAAGGCTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAAGATGATGA 1965

Qy   1242 ATTAAGCCGGCACCTATTGGAGGCACGACTTTCAGAGAAGCAGAACGACGATGT 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1966 ATTAAGCCGGCACCTATTGGAGGCACGACTTTCAGAGAAGCAGAACGACGATGT 2025

Qy   1302 ACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2026 ACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2085

Qy   1362 GACTGTACGAATATTTCTGACAGAGACCAGCCTTGGAAGGACTAGAGAAACTCTACCAGGA 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2086 GACTGTACGAATATTTCTGACAGAGACCAGCCTTGGAAGGACTAGAGAAACTCTACCAGGA 2145

Qy   1422 GCCCAGAGAGCTGCCTCCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGGAAAGCA 1481
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2146 GCCCAGAGAGCTGCCTCCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 2205

Qy   1482 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCCGCTGACTGGCAGAG 1541
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2206 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCCGCTGACTGGCAGAG 2265

Qy   1542 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGAGGCCACGGATGAGCTGGA 1601
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2266 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGAGGCCACGGATGAGCTGGA 2325

Qy   1602 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGGGATCTCCT 1661
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2326 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGGGATCTCCT 2385

Qy   1662 CATTGACTTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAATTGCGCC 1721
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2386 CATTGACTCTCTCCAAGATCACCTCCAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 2445

Qy   1722 TCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCACTTTGGGCAT 1781
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2446 TCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCACTTTGGGCAT 2505
```

```
Db   8620 TGACCAGTGGAAGCGGCTGCCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA 8679

Qy   1233 AGATGATGAATTAAGCCGGCACCTATTGGAGGCACGACTTTCAGAGAAGCA 1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8680 AGATGATGAATTAAGCCGGCACCTATTTCGACGAATATTTCGACGAGGCACTAGAGAAGCA 8739

Qy   1293 GAACGAGTGCATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8740 GAACGAGTGCATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 8799

Qy   1353 TACTCTTGAGACTCTGACGAATATTTCTGACAGAGACTAGAGAGACTAGAGAAACT 1412
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8800 TACTCTTGAGACTGTGACGAATATTTCTGACAGAGACTAGAGAGACTAGAGAAACT 8859

Qy   1413 CTACCAGGAGCCCAGAGAGCTGCCTCCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCT 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8860 CTACCAGGAGCCCAGAGAGCTGCCTCCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCT 8919

Qy   1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCCACTCCGCTGA 1532
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8920 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCCACTCCGCTGA 8979

Qy   1533 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8980 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGGA 9039

Qy   1593 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGCCGTGGG 1652
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9040 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGCCGTGGG 9099

Qy   1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9100 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 9159

Qy   1713 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCAC 1772
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9160 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCAC 9219

Qy   1773 TTTGGGATTCAGCTTCACCGTATAACCTCAGACACTCTGAAGAACCTGAACACCAGATG 1832
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9220 TTTGGGATTCAGCTTCACCGTATAACCTCTGGAAGAACCTGAACACCAGATG 9279

Qy   1833 GAAGCTTCTGCAGGTTGGCCGTCGGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA 1892
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9280 GAAGCTTCTGCAGGTTGGCCGTCGGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA 9339

Qy   1893 CTTTGGTGGTCCAGCATCTCAGCACTTTCTTCTTTGTCCACGTGTCCTGTCTGCAGGGTCCCTGGGAGGAGAGC 1952
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9340 CTTTGGTGGTCCAGCATCTCAGCACTTTCTTCTTTGTCCACGTGTCCTGTCTGCAGGGTCCCTGGGAGGAGAGC 9399

Qy   1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 9448
```

```
Qy  1782  TCAGTCTCACGCGTATAACCTCAGCACTCTGGAAGACCTGAAGACACCAGATGGAAGCTTCT  1841
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2506  TCAGTCTCACGCGTATAACCTCAGCACTCTGGAAGACCTGAAGACACCAGATGGAAGCTTCT  2565

Qy  1842  GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTCGCATGAAGCCCACCAGGGACTTTGGTCC  1901
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2566  GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTCGCATGAAGCCCACCAGGGACTTTGGTCC  2625

Qy  1902  AGCATCTCAGCACTTCTTCTCCACGTCGTCGTCCAGGGTCCCTGGGAGAGAGGCCATCTCGCC  1961
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2626  AGCATCTCAGCACTTCTTCTCCACGTCGTCGTCCAGGGTCCCTGGGAGAGAGGCCATCTCGCC  2685

Qy  1962  AAACAAAGTGCCTACTATATCAACCACACGAGACTCAAACA  2001
          |||||||||||||||||||||||||||||||||||||||||
Db  2686  AAACAAAGTGCCCCACTACTATATCAACCACACGAGACTCAAACA  2725
```

RESULT 13
AF070485
LOCUS   AF070485   13887 bp   mRNA   linear   MAM 09-DEC-1998
DEFINITION Canis familiaris dystrophin mRNA, complete cds.
ACCESSION   AF070485
VERSION   AF070485.1   GI:3982750
KEYWORDS   .
SOURCE   Canis familiaris (dog)
ORGANISM   Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1 (bases 1 to 13887)
AUTHORS   Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
TITLE   Direct Submission
JOURNAL   Submitted (04-JUN-1998) ANRI, Pathology, University of Western
            Australia, Verdun Street, Nedlands, WA 6018, Australia
FEATURES   Location/Qualifiers
     source   1..13887
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /strain="breed Golden retriever"
                     /db_xref="taxon:9615"
     5'UTR   <1..289
     CDS   290..11332
                     /codon_start=1
                     /product="dystrophin"
                     /protein_id="AAC3646.1"
                     /db_xref="GI:3982751"
                     /translation="MLWWEWEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNVNVNKALRVLQKNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIAKYQLGIEKLLDPEDV
                     ATTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRFPSQVTREHFQIHHQMHYSQ
                     QITVSLAQGYERAPSFPKPRFKSYAYTQAAYVTTSDPTRSPLPSQHLETPBDKSFGRS
                     LTETEANLDSYGTALEEVLSWLLSAEDALQAQGEISNDVEEVKEQFHTHEGYMMDLTS
                     HQGRVGNVLQLGSQLIGTGKLSEDEETEVQEQMNLLNSRWEBCLRVASMEKQSNLHKVL
                     MDIQNQQLKELNDWLTKTEERTRKMEEKPLGPDIEDLKRQVQQHKVLQEBDLEQQVRV
                     NSLTHMVVVDBSSGDHATAALEBQLKVLGDRWANICRWTEDRWVLLQDILLKWQRFT
                     EEQCLFSAWLSEKEDAVNKIHTTGFKDQSEVLSNLQKLAVLKTDLEKKKQTMDKLCSL
                     NQDLLSALKNTVAHKMEAWLIDPPPPQKRQIIVDSEIRKRLDVDITRLHSWITRSE
                     METVTMVTTREHILVHAQEELPPPQWKKPSDLKEKVNAIERRKAEKFRKLQDASRSAQALVEQMVNEGVN
                     AVLQSPEFAIYRKEGNPSDLKEKVNAIERRKAEKFRKLQDASRSAQALVEQMVNEGVN
                     ADSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIITFYNQLQOIEKQMTTAENWLKT
                     QPTTSEPTAIKSQLKICKDEINRLSALQPQIERLKIQSIALKEKGQGPMLDADFVA
                     FTNHFNQVFADVQAARELQLTIFDSLPPMRYQETMSTILTWIQQSETKLSIPQVTVTE
                     YDIMEQRLGELGALQSSLQEQGNGLNYLSFTVELGR
                     WKKLSSQLVEHCQKLESQMAKLRKIQNHIKTLKKWITVPFLKKWITEVDVFLKBEWPALGDSEILKR
                     QLKQCRLLVNDIQTIQPSLNSVNEGAQKMKNEAEPFAGRLETELRELNTWDYMCRQ
                     VYARKEALKGGLDKTVSLQKDLSEMHEWMTQAEBEYLERDFEYKTPDELQTAEWMKR
                     AKEBAQQKEAKVKLITSVNSVIAQAPPAAQEALKELDTIITNYQWLCTRLNGKCKT
                     LEEVWACWHELLSYLEKANKMLSEVEVKLKTENISGGAEBIAEVLDSLENLMQHSED
                     NPNQIRILAQTLTDGGVMDELINEELETFNSRWRELHEEAVRRQKLLEQSICSAQEIE
                     KSLHLIQESLSSIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLSEMKKHNQGK
                     ETAQRVLSQIDVAQKLQDVSMKFRLFQKPANFEQRLQBSSMILDEVKOMHLPALETKS
                     VEQEVVQSQLNHCVNLYKSLSEVKSEVEMVIKTGRQIVQKKQTENPKELDERVTALKL"

```
Qy  813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTCTTGCCTG  872
         ||||||||||||||||||||||| ||||| ||| ||||||||||||||||||||||||||
Db 8326  GGAAGAAACCCACAGATTACTGCAACAGTTCCCCTTGGACCTGGAAGAGTTCCTTGCCTG  8385

Qy  873  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGCTACCCGTAAGGAAAGGCT  932
         ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db 8386  GCTTACAGAAGCTGCAAACAACTGCCAACGTCCTGCAGGATGCCACCCATAAGGAAAGGCT  8445

Qy  933  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA  992
         |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db 8446  TCTAGAAGATTCCAAGGAGTAAGAGAGCTGATGAAACAATGGCAAGACCTCCAAGGAGA  8505

Qy  993  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAATCCTGAG  1052
         ||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||
Db 8506  AATCGAAGCTCACACAGATATCTATCACAACCTGGACGAGAAAATGGCCAAAAAGTCCTGAG  8565

Qy 1053  ATCCCTGGAAGGTTCCGATGATGCAGTCGTTGTTACAAAGACGTTTGGACATAACAGAACTT  1112
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db 8566  ATCCCTGGAAGGTTCTGACGATGCCAGCGTTGTTGCAAAGACGTTTGGACATAACAGAACTT  8625

Qy 1113  CAAGTGGAGTGAACTTCCGGAAAAAAGTGTCTCCAACATTAGGTCCCATTTGGAAGCCAGTT  1172
         |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db 8626  CAAGTGGAGCGAACTTCCGGAAAAAAGTGTCTCAACATTAGGTCCCACTGGAAGCCAGTTT  8685

Qy 1173  TGACCAGTGGAAGCGTCGCACCTTCTGTGGTGGCTCACAGTGAA  1232
         ||||||||||||||||||||||||||||||||||||||||||||
```

HYNELGAKVTERKQQLEKCLKLSRKMREKMNALTEWLAATDMELTKRSAVEGMPSNLD
SEVAWGKATQKEIBKQKVHLKSVTBVGEALKTVGKKEMLVEDKLSLLNSNWIAVTSR
ABEWINLLLEYQKMETFDQNVDYITNWIIQADALLDESEKKPQQKEDILKRLKAEM
NDIRPKVDSTRDQAANLMANRGDHCRKVVEPKISELNHRFAAISHRIKTGKASIPLKE
LEQFNSDIQKLLEPLBAEIQQGVNLKEEDPNKDMSEDNEGTVKELLQRGDNLQQRITD
ERKRBEIKIKQQLLQTKHNALKDLRSQRRKKALEISHQWYQYKRQADDLLKCLDDIEK
KLASLPEPRDERKISIDRELQKKKEELNAVRRQABGLSEDGAMAVEPTQIQLSKRW
REIESKFAQFRRLNFAQITHVEBSVVAMTEDMPLEISYVPSTYLTEITHVSQALSEV
BELLNAPDLCAQDFEDLFKQEBSLRNIKDSLQQISGRIDIIHNKKTAALHSATPAERA
KLQEALSRLDFQWERVNNMYKDRQGRFDRSVEKWRRFHYDMKILNQWLTEAEQFLKKT
QIPENWEHAKYKWYLKELQDGIGQRQSVVRVLNATGEBLIIQQSKTDASILQEKLGSL
NLRWQEVCKQLAERKKRLEEQKNIISEFQRDVNEFVLWEBADNVANIPLEPGNEQQL
KEKLEQVKLLAEELPLRQGILKQLNETGGTVLVSAPLSPEBQDKLENKLKQTNLQWIK
VSRNLPEKQEIEAHVKDLGQLEBQLNHLLLWLSPIRNQLEIYNQPNQTGPFDIKEIE
VAVQAKQPDVEGILSKGQHLYKERPATQPAKRKLEDLSSDWKVTQLIQELRAKQPGP
APGLTTVRAPPSQTVLTVTQPAVTKETAISKLEMPSSLLLEVPALADFNRAWTELTDW
LSLLDRVIKSQRVMVGDLEDINEMIIKQKATLQDLEQRPQLEELITAAQNLKNKTSN
QEARTIITDRIERICQSWDEVQEHLQNRRLQLTEMLKDSTOWLEAKEBAEQVLGQARA
KLESWKEAPYTVDAIQKKITBTKQLAKDLRQWQINVDVANDLALKLLRDYSADDTRKV
HMITENINAGWASIHKRLSEREEAALEETHRLLQQFPLDLEKFLAWLTBAETTANVLQD
ATHKRLLEDSKGVRELMKQWDLQGEIEAHTDIYHNLDENGQKVLRSLEGSDDAALL
QRRLDMNNFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELIVWLQLKDDELSRQAP
IGGDFPAVQKQNDVHRAFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLYQEPREL
PBEERAQNVTRLLRKQABEVNTQWEKLNVHSADWQRKIDEALERLAQELQEATDELDLK
LRQAEVIKGSWQPVGDLLLIDSLQDHLEKVKALRGEITPLKENVSYVNDLARQLTTLGI
QLSPYNLNTLEDLNTRWKLLQVAIEDRIRQLHEAHRDFGPASQHFLSTSVQGPWERAI
SPNKVPYYINHBTQTTCWDHPKMTBLYQSLADLNNVRFSAYRTAMKLRLYQKALCLDL
LSLSAACDALDQHNLKQNDQPMDILQVINCLTTIYDRLEQEHNNLVNVPLCVDMCLNW
LLNVFDTGRTGRIRVLSFKTGIISLCKAHLEBDKYRYLFKQVASSTGFCDQRRLGLLH
DSIQIPRQLGEVASFGSNIEPSVRSCFQFANNKPEIEAALFLDWMRLEPQSMWVLKPY
LHRVAAAETAKHQAKCNICKECPIIGFRYRSLKFNYDICQSCFFSGRVAKGHKMHYP
MVEYCPTTSGEDVRDFAKVLKNKFRTKYFAKHPRMGYLPVQTVLEGDNMETPVTLI
NFWPVDSAPASPQLSHDDTHSRIEHYASRLKKMENSNGSYLNDSISPNSIDEHLL
IQHYRVSLNQBSPLSQPRSPAQIILSLESEBERGELERIIADLEGRNRNLQAEYDRLKQ
QEHKGLSPIFSPPEMMPTSPQSRDAEBIAEAKLLRQHKGRLEARMQILEDHNKQLE
SQLHRLRQLLEBQPQABAKVNGTTVVSSPTSLQRSDSSQPMLLRVVGSQTSESMGEEDL
LSPPQDTSTGLEEVMEQLNHSFPSSRGRNTPGKPMREDTM"

3'UTR   11333..>13887
repeat_region   11789..11828
                /rpt_type=tandem
                /rpt_unit="ca"

ORIGIN

Query Match   51.5%;   Score 1030.6;   DB 4;   Length 13887;
Best Local Similarity   91.7%;   Pred. No. 1.7e-247;
Matches 1090;   Conservative   0;   Mismatches   99;   Indels   0;   Gaps   0;

```
                                  /organism="Mus musculus"
                                  /mol_type="unassigned DNA"
                                  /db_xref="taxon:10090"

ORIGIN

     Query Match       50.2%;   Score 1004;   DB 6;   Length 13815;
     Best Local Similarity   90.3%;   Pred. No. 8.2e-241;
     Matches 1073;   Conservative  0;   Mismatches 115;   Indels  0;   Gaps  0;

Qy    813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGACCAGTTCTTGCCTG    872
          ||||  |||||||||||||||||| |||| ||||||||||| |||| ||||||
Db   8251 GGAAGAAACTCATAGATTACTGCGACCAGTTCCCTGACCTGGAGAAGTTTCTTTCCTG   8310

Qy    873 GCTTACAGAGCTGAAACACTGCCAATGTCCAGAAGTCCAGAAGGCT    932
          ||||||||||||||||||||||||||||  |||||||||||||||||
Db   8311 GATTACGGAGCAGAAACACTGCCAATGTCCTACAGGACGCTTCCCGTAAGGAAGAGAAGCT   8370

Qy    933 CCTGAAAGACTCCAAGGAGGAGTAAAAGAGTGATGAAACAATGGCAAGACCTCCAAGGTGA    992
          |||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db   8371 CCTAGAAGACTCCAAGGGGAGTCAGAGACTGATGAAACCATGGCAAGATCTCCAAGGAGA   8430

Qy    993 AATTGAAGTCACACAGAGTGTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAG   1052
          ||||||||||||||||||||||| ||||||||||||||||||||||| |||||||||||||
Db   8431 AATTGAAACTCACACAGATATCATCACAATCTTGATGAAAATCACAAATCCTGAG   8490

Qy   1053 ATCCCTGGAAGGTTCGATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTT   1112
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8491 ATCCCTGGAAGGTTCGGATGATGAAGCACACCCCTGTTACAAAGACGTTTGGATAACATGAATTT   8550

Qy   1113 CAAGTGGAGTGAACTTCGAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTC   1172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8551 CAAGTGGAGTGAACTTCAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCAAGTTC   8610

Qy   1173 TGACCAGTGGAAGCGTCTGCCACCTTTCTTCTTCAGGAACTTCTCGGTATGGCTCGTACAGCTGAA   1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8611 TGACCAGTGGAAGCCGTTTGCATCTTTCTTCTTCAGGAACTTCTGTTTCGTGTTGGCTACAGCTGAA   8670

Qy   1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGACACTTTCCAGCAGTTCAGAAGCA   1292
          ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db   8671 AGATGATGAGTTAAGCCGGCAGGCACCCATTGGAGGGACACTTTCCAGCAGTTCAGAAGCA   8730

Qy   1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG   1352
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8731 GAATGATGTACATAGGGCCTCGAGGAGCTGCAGAGCTGCATGAGAGCCCACAGGGA   8790

Qy   1353 TACTCTTGAGACTGTACGAATATTTCGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT   1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8791 TACTCTTGAGACTGTGAGAGTGTGAGACTGTGAGATATATTTCGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT   8850

Qy   1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAAGAGCCCAGATGTCACTCGGCTTCT   1472
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8851 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAAGAGAAAGAGTCAGATGTCACTCGGCTCCT   8910

Qy   1473 ACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCCAGTGA   1532
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8911 ACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGAAATTGAACCTGCGCTCAGTGA   8970

Qy   1533 CTGGCAGAGAAAAATAGAGACCCTTGAAAGACTTCAAGAGGCCACGGA   1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8971 TTGGCAGAGAAAAATAGATGAAGCTTCCTTGAAAGACTTCCAGGAAGTGCCGA   9030

Qy   1593 TGAGCTGGACTTCAAGCTGAGGTGATCGAGGATCAAGGATCCTGGCAGCCAGTTC   1652
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9031 TGAACTGGACTTCAAGTTGCGCAAGCTGAGGTGATCAAGGATCCTGGCAGCCAGTTGGG   9090

Qy   1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTTCGAGAAAGTCAGGACACTTCGAGGAGA   1712
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9091 TGACCTCCTCATTGACTCTCTCCAAGATCACCTCCGAGAAAGTCAGGACGCTTCGAGGAGA   9150

Qy   1713 AATTGCGCCCTGAAAGAGAACGTGAGCCCAGTGACCTTGCTCTGCCAGTTACCAC   1772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9151 AATTACACCTCTGAAAAGTGCCTCAATGTCAGTCAGTCACTGACCTTGCTCTGCCAACTCACTAC   9210
```

```
b   8686 TGACCAGTGGAAGCGTCTGCCACCTTTCTTCTTCAGGAACTTCTGGTATGGCTCCAGCTGAA   8745
y   1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGACACTTTCCAGCAGTTCAGAAGCA   1292
b   8746 AGAATGATGAGTTAAGCCGGCAGGCACCCATTGGAGGGACACTTTCCAGCGGGTGCAGAAGCA   8805
y   1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG   1352
b   8806 GAATGATGTACACAGGGCCTTCAAGAGGAATTGAAAACGAAAGAACCTGTAATCATGAG   8865
y   1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT   1412
b   8866 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT   8925
y   1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAAGAGCCCAGATGTCACTCGGCTTCT   1472
b   8926 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAAGAGAATGTCACACGGCCTCCT   8985
y   1473 ACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCCAGTGA   1532
b   8986 ACGAAAGCAGGCTGAGGAGTCAACTCAGTGGGAAAAACTGAACGTGCCACTCTGCAGA   9045
y   1533 CTGGCAGAGAAAAATAGAGACCCTTGAAAGACTTCAAGAGGCCACGGA   1592
b   9046 CTGGCAGAGAAAAATAGACGAAGGGCCCTCGAGGAGCTTCAGGAGGCAACAGA   9105
y   1593 TGAGCTGGACTTCAAGCTGAGGTGATCGAGGATCAAGGATCCTGGCAGCCAGTTC   1652
b   9106 TGAGCTGGATCTCAAACTACGTCAGGCAGAGGATCAAGGATCCTGGCAGCCAGTTGGG   9165
y   1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTTCGAGAAAGTCAGGACACTTCGAGGAGA   1712
b   9166 TGACCTCCTCATTGACTCTCTCCAAGATCACCTCCGAGAAAGTCAGGCGCGCTTCGAGGAGA   9225
y   1713 AATTGCGCCCTGAAAGAGAACGTGAGCCCAGTGACCTTGCTCTGCCAGTTACCAC   1772
b   9226 AATTACACCTCTGAAAAGTGCCTCAATGTCAGTCACGTCACCTTGCTCTGCCAACTCACTAC   9285
y   1773 TTTGGGCATTCAGCTCTCACCGTATAACCTCCAGCACTCTGGAAGACTCTGAAACACCAGATG   1832
b   9286 GTTGGGCATTCAGCTGTCACCATATAACCTCCAACACTCTGGAGAGACCTGAAACACCAGATG   9345
y   1833 GAAGCTTCTGCGAGGTGGCGCGTTCAGGAGGACCTGCGAGGAGCTGCATGAGAGCCCACAGGGA   1892
b   9346 GAAGCTTCTGCGAGGTGGCCATTGAGGACCTGCATTCCACTTTCTTCCACTTCCTGGGAGACCGCACAGGA   9405
y   1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTGTCCAGGGTCCTGGGAGGAGAGC   1952
b   9406 CTTTGGATCCAGCCGCTCCCAGCACTTTCCACTTTCGTGTCCAGGGTCCTGGGAGGAGGC   9465
y   1953 CATCTGCCAAACAAAGTGCCCTACTACTATATCAACCACGAGACTCAAACA   2001
b   9466 CATCTCACCAAACAAAGTGCCCTACTATATCAACCACCAGAGACCCAAACA   9514
```

```
RESULT 14
X306153
LOCUS       AX306153           13815 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 904 from Patent WO0188188.
ACCESSION   AX306153
VERSION     AX306153.1  GI:17645441
KEYWORDS
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
  TITLE     Method for examining ischemic conditions
  JOURNAL   Patent: WO 0188188-A 904 22-NOV-2001;
            School Juridical Person Nihon University (JP)
FEATURES             Location/Qualifiers
     source          1..13815
```

```
Qy   1773  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGACACCAGATG  1832
           |||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db   9211  ACTGGGCATTCAGCTCTCACCTTATAACCTCAGCACTTGGAAGATCTGAATACCAGATG  9270

Qy   1833  GAAGCTTCTGCGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCAGGGA  1892
           |||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||
Db   9271  GAGGCTTCTTACAGGTGGCTGGTGGAGGACCGTGTCAGACAGCTGCATGAAGCCCAGGGA  9330

Qy   1893  CTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC  1952
           ||||| |||||||||||||||||| |||| |||||||||||||||||||||||||||
Db   9331  CTTTGGTCCTGCATCCCAGCACTTCCTTCCACTTCAGTTCAGGGTCCCTGGGAGAGAGC  9390

Qy   1953  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC  2000
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   9391  CATCTCACCAAACAAAGTGCCCTACTATATCAACCACGAGACCCAAAC  9438
```

RESULT 15

AX538582

```
LOCUS       AX538582            13815 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 2 from Patent WO0229956.
ACCESSION   AX538582
VERSION     AX538582.1  GI:25271088
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 2 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..13815
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN
```

```
Query Match         50.2%;  Score 1004;  DB 6;  Length 13815;
Best Local Similarity   90.3%;  Pred. No. 8.2e-241;
Matches 1073;  Conservative    0;  Mismatches 115;  Indels    0;  Gaps    0;
```

```
Qy    813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG  872
           |||||||||||||||||||||||||| |||||||||| |||||||||||||||||||||
Db   8251  GGAAGAAACTCATAGATTACTGCGCAGCAGTTCCCCTGGACCTGGAGAAGTTTCTTCCTG  8310

Qy    873  GCTTACAGAGACTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT  932
           |||||||||||||||| ||||||||||||||| |||||||| ||||||||||||||||||
Db   8311  GATTACGGAAGACTCCAGAAGCAGAAACAACTGCCAATGTCCTACAGGACGCGCTCCCGTAAGGAGAGAGCT  8370

Qy    933  CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACTCCAAGGTGA  992
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   8371  CCTAGAAGACTCCAGGGAGTCAGGAGGTCAGGAGCTGATGAAACATGGCAAGATCTCCAAGGAGA  8430

Qy    993  AATTGAAGCTCACACAGATGTTATCACAACCTGGATGATGAAAACAGCCAAAAAATCCTGAG  1052
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8431  AATTGAAACTCACACAGATATCTCATCACAACCTGGATGATGAAAATGGCCAAAAAATCCTGAG  8490

Qy   1053  ATCCCTGGAAGGTTCCGATGATGCAGTCGAGTCCTGGAAAAAGTTTTGATAACATGAACTT  1112
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8491  ATCCCTGGAAGGTTCGGATGATGCAGTCGAGTCCTGTTACAAAGACGTTTGATAACATGAATTT  8550

Qy   1113  CAAGTGGAGTGAAGCTTCGAGACTTCAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  1172
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8551  CAAGTGGAGTGAAGCTTCGAGACTTCAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCAAGTTC  8610

Qy   1173  TGACCAGTGGAGGGTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA  1232
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8611  TGACCAGTGGAGGGTCTGCATCTTCTCTGCAGGAACTTCTTGTTTGGCTACAGCTGAA  8670

Qy   1233  AGATGATGAATTAAGCCGGCAGCCGACTTTCCAGCAGTTCAGAAGCA  1292
```

```
Search completed: April  5, 2004, 02:39:39
Job time : 5220.76 secs
```

```
Db   8671  AGATGATGAACTGAGCCGGTCAGGCAGCCATCGGTGGTGATTTCCCAGCAGTTCAGAAGCA  8730

Qy   1293  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
           |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8731  GAATGATATACATAGGGCCTTCAAGAGAGGAATTGAAAACTAAAGAACCTGTAATCATGAG  8790

Qy   1353  TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGACAGAGGACTAGAGAAACT  1412
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   8791  TACTCTGGAGACTGTGAGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACT  8850

Qy   1413  CTACCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCT  1472
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   8851  CTACCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAAGAGCTCAGAATGTCACTCGGCTCCT  8910

Qy   1473  ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGAACCTGACCTGA  1532
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8911  ACGAAAGCAGGCTGAGGAGGTCAACGCTGAATGCGAATGACAAATTGAACCTGAACCTGACCTGA  8970

Qy   1533  CTGGCAGAGAAAATAGGAGACCCTTGAAAGACTCCAGGAGACTTCAAGAGGCCACGGA  1592
           ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db   8971  TTGGCAGAGAAAAATAGGAGACTCTTGAAAGACTCCAGGAACTTCAGGAAGCTGCCGA  9030

Qy   1593  TGAGCTGGACTCAAGCTGCGCCAAGCTGACGTGATCAGGGTCGAGCCCCGTGGG  1652
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9031  TGAACTGGACCTCAAGTCGCGCCAAGCTGAGTGATCAAGGATCCTGGCCGCCAGTGGGG  9090

Qy   1653  CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCACCTCAAGGAGACACTCAGGAGATG  1712
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   9091  GGATCTCCTCATTGACTCTCTCTGCAAGATCACCTTATAAAGTCAAGGCACTTCGGGGAGA  9150

Qy   1713  AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC  1772
           ||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db   9151  AATTGCACCTCTTAAAGAAGTGTCAATCGTGTCAATGACCTTGCACATCAGCTGACCAC  9210
```

```
Copyright        GenCore version 5.1.6
                 Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

run on:     April  4, 2004, 16:52:43 ; Search time 3673.37 Seconds
                                        (without alignments)
                                        17079.796 Million cell updates/sec

title:           US-09-845-416-2_COPY_900_3000
perfect score:   2101
sequence:        1 gagctatgcctacacacagg........acctcagcactctggaagac 2101

scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

searched:        27513289 seqs, 1493109276 residues

total number of hits satisfying chosen parameters:   55026578

minimum DB seq length: 0
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

database :       EST:*
                  1:   em_estba:*
                  2:   em_esthum:*
                  3:   em_estin:*
                  4:   em_estmu:*
                  5:   em_estov:*
                  6:   em_estpl:*
                  7:   em_estro:*
                  8:   em_htc:*
                  9:   gb_est1:*
                 10:   gb_est2:*
                 11:   gb_htc:*
                 12:   gb_est3:*
                 13:   gb_est4:*
                 14:   gb_est5:*
                 15:   em_estfun:*
                 16:   em_estom:*
                 17:   em_gss_hum:*
                 18:   em_gss_inv:*
                 19:   em_gss_pln:*
                 20:   em_gss_vrt:*
                 21:   em_gss_fun:*
                 22:   em_gss_mam:*
                 23:   em_gss_mus:*
                 24:   em_gss_pro:*
                 25:   em_gss_rod:*
                 26:   em_gss_phg:*
                 27:   em_gss_vrl:*
                 28:   gb_gss1:*
                 29:   gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1103.4	52.5	3870	11	BC036103	BC036103 Homo sapi
2	866.6	41.2	3056	11	AK044536	AK044536 Mus muscu
3	846.2	40.3	5691	29	AY399453	AY399453 Homo sapi
4	633.4	30.1	5697	29	AY399455	AY399455 Mus muscu

No.	c	Score	%	Length	DB	ID	Description
5		597.2	28.4	728	14	CB228986	CB228986 AGENCOURT
6		592.4	28.2	834	12	BI729851	BI729851 603349511
7		514.4	24.5	5676	29	AY399454	AY399454 Pan trogl
8	c	509	24.2	595	14	CB177816	CB177816 is21c01.x
9		493.2	23.5	1047	14	CB850319	CB850319 MRA-0070
10		491.6	23.4	750	12	BI730168	BI730168_603349711
11		486.6	23.2	579	9	AL121550	AL121550 DKFZp762L
12		424.2	20.2	663	12	BM488464	BM488464 pgm2n.pk0
13		389	18.5	644	13	BU313510	BU313510 603540290
14		343	16.3	2334	11	BC011062	BC011062 Mus muscu
15		334.6	15.9	9915	29	AY407022	AY407022 Homo sapi
16	c	324.6	15.4	402	14	CB547284	CB547284 AMGNNUC:S
17		312.2	14.9	9691	29	AY407024	AY407024 Mus muscu
18		311.8	14.8	9096	29	AY407023	AY407023 Pan trogl
19		275.2	13.1	1122	14	CF109978	CF109978 Shultzomi
20		265.8	12.7	772	12	BI250598	BI250598 602993659
21		226	10.8	250	10	BF963618	BF963618 QV2-NN004
22		217.2	10.3	851	13	BU201022	BU201022 603952191
23		213.2	10.1	2874	29	AY408546	AY408546 Homo sapi
24		212.2	10.1	3051	11	BC036095	BC036095 Homo sapi
25		203.6	9.7	1541	11	AK034383	AK034383 Mus muscu
26		203.6	9.7	2874	29	AY408548	AY408548 Mus muscu
27		192	9.1	3753	11	AK081426	AK081426 Mus muscu
28		192	9.1	778	14	CB524596	CB524596 UI-M-FY0-
29		184.4	8.8	812	14	CF748008	CF748008 UI-M-HE0-
30	c	183	8.7	763	14	CB518960	CB518960 UI-M-GH0-
31	c	182.4	8.7	645	10	AW467977	AW467977 he29g08.x
32		177	8.4	646	12	BI289102	BI289102 UI-R-DK0-
33		172.8	8.2	835	12	BI553820	BI553820 603190772
34		172.8	8.2	2874	29	AY408547	AY408547 Pan trogl
35		166.2	7.9	907	13	BU122401	BU122401 603003073
36		165.2	7.9	600	12	BI988528	BI988528 4012-24 M
37		162	7.7	427	10	BE817979	BE817979 CM2-BN027
38	c	158.2	7.5	569	29	CE228097	CE228097 tigr-gss-
39	c	158.2	7.5	615	14	CA377239	CA377239 655768 NC
40		153.2	7.3	575	29	CG519934	CG519934 OST84444
41		153.2	7.3	885	11	AK020881	AK020881 Mus muscu
42	c	150.6	7.2	515	28	AZ780914	AZ780914 2M0018O10
43		150.2	7.1	497	29	CG525393	CG525393 OST100373
44		150.2	7.1	497	29	CG653669	CG653669 OST419391
45		149.2	7.1	497	29	CG553287	CG553287 OST165637

ALIGNMENTS

```
RESULT 1
BC036103
LOCUS       BC036103                3870 bp    mRNA    linear   HTC 19-NOV-2003
DEFINITION  Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker
            types), mRNA (cDNA clone IMAGE:5274415), with apparent retained
            intron.
ACCESSION   BC036103
VERSION     BC036103.1  GI:23271310
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3870)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
```

TITLE   Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
         Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
         Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
         Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
         Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
         human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED   12477932
REFERENCE 2 (bases 1 to 3870)
AUTHORS Strausberg,R.
TITLE   Direct Submission
JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian
         Gene Collection (MGC), Cancer Genomics Office, National Cancer
         Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
         USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
         Email: cgapbs-r@mail.nih.gov
         Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
         cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
         Toshiyuki and Piero Carninci (RIKEN)
         cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Sequencing Group at the Stanford Human Genome
         Center, Stanford University School of Medicine, Stanford, CA 94305
         Web site: http://www-shgc.stanford.edu
         Contact: (Dickson, Mark) mcd@paxil.stanford.edu
         Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
         R. M.

         Clone distribution: MGC clone distribution information can be found
         through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
         Series: IRAK Plate: 48 Row: f Column: 7
         This clone was selected for full length sequencing because it
         passed the following selection criteria: similarity but not
         identity to protein.
         This clone has the following problem: retained intron.
FEATURES   Location/Qualifiers
     source   1..3870
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:9606"
                     /clone="IMAGE:5274415"
                     /tissue_type="Brain, hippocampus"
                     /clone_lib="NIH_MGC_95"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
ORIGIN

Query Match   52.5%; Score 1103.4; DB 11; Length 3870;
Best Local Similarity 99.9%; Pred. No. 8.2e-235;
Matches 1104; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGACCCTACACGAGCCCATT 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGACCCTACACGAGCCCATT 1102

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1162

Qy   121 TGAAGTAAACCTGGACCCGTTATCAAACAGTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 TGAAGTAAACCTGGACCCGTTATCAAACAGTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1222

Qy   181 TGCTGAGGACACATTGGAAGCACAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 TGCTGAGGACACATTGGAAGCACAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1282

Qy   241 CCAGTTCATCATCATGAGGGGTACAGTGATGGAGATTTGACAGCCATCAGGGGCGGGGTTGG 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1283 CCAGTTCATCATCATGAGGGGTACAGTGATGGAGATTTGACAGCCCCATCAGGGGCCGGGGTTGG 1342

Qy   301 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1343 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1402

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1403 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1462

Qy   421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1463 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1522

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAGAAAATGAGGAGAAGAACAAGGAGGAGA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1523 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAGAAAATGGAGAAAATGGGGAGAAGA 1582

Qy   541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAACACATAAGGTGCTTCA 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1583 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACAACACATAAGGTGCTTCA 1642

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAATTCTCCTCACTCCACATGGTGGTGGTAGT 660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1643 AGAAGATCTAGAACAAGAACAAGTCAATTCTCCTCACTCCACATGGTGGTGGTAGT 1702

Qy   661 TGATGAATCAGTGGAGATCACGCAACTGCTGTTGGAAGAACAACTTAAGGTATTGGG 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1703 TGATGAATCAGTGGAGATCACGCAACTGCTGTTGGAAGAACAACTTAAGGTATTGGG 1762

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTCGGGTCTTTTACAAGACAT 780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1763 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTCGGGTCTTTTTACAAGACAT 1822

Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1823 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1882

Qy   841 AAAGAAGATGAGATCCAGTGAACAAGATCCACACAACTGGCTTAAAGATCAAAATGAAATGTT 900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1883 AAAGAAGATGAGATCCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1942

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCGGATCAGAAAAAGAAAAAGCAATCCAT 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1943 ATCAAGTCTTCAAAAACTGGCCGCGTTTTAAAAGCGGATCAGAAAAAGAAAAAGCAATCCAT 2002

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2003 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2062

Qy   1021 CCAGAAGACGGAAGCATGGCCTGGATAACTTTGCCCGGTGTGTTGGGATAATAATTTAGTCCAAAA 1080
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2063 CCAGAAGACGGAAGCATGGCCTGGATAACTTTGCCCGGTGTGTTGGGATAATAATTTAGTCCAAAA 2122

Qy   1081 ACTTGAAAAGAGTACAGCACAGACT 1105
             |||||||||||||||||||||||||
Db   2123 ACTTGAAAAGAGTACAGCACAGATT 2147

RESULT 2
AK044536
LOCUS   AK044536   3056 bp   mRNA   linear   HTC 20-SEP-2003
DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library,
            clone:A930019F21 product:dystrophin, muscular dystrophy, full
            insert sequence.
ACCESSION   AK044536
VERSION   AK044536.1 GI:26090404
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.

TITLE   High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
 PUBMED   10349636
REFERENCE 2
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
          Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
 PUBMED   11042159
REFERENCE 3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
          Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
          Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
          Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
          Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
          Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
          Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system-384-format
          sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
 PUBMED   11076861
REFERENCE 4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
          FANTOM Consortium.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3056)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
          Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
          Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
          Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
          Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
          Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
          Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
          Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
          Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
          Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
          Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
          Physical and Chemical Research (RIKEN), Laboratory for Genome
          Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
          RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
          Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
          URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
          Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Retina RNA was provided by Dr. Stefano Gustincich (Department of
          Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
          02115, USA) whose assistance is gratefully acknowledged. Please
          visit our web site for further details.
          URL:http://genome.gsc.riken.go.jp/
          URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..3056
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"

```
                     /db_xref="FANTOM_DB:A930019F21"
                     /db_xref="MGI:2409854"
                     /db_xref="taxon:10090"
                     /clone="A930019F21"
                     /tissue_type="retina"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    1..3056
                     /note="dystrophin, muscular dystrophy (MGD|MGI:94909,
                     GB|NM_007868, evidence: BLASTN, 100%, match=99)"

ORIGIN

Query Match             41.2%;  Score 866.6;  DB 11;  Length 3056;
Best Local Similarity   86.5%;  Pred. No. 4.2e-182;
Matches  956;  Conservative    0;  Mismatches  149;  Indels    0;  Gaps    0;

Qy     1 GAGCTATGGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACCGGAGCCCATT   60
           |||| |||| |||||||||||||||||||| | ||||||||||  |||||   |||||
Db   944 GAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACCTCTGATTCCACACAGCAGCCCCTA 1003

Qy    61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCAGTTCATTGATGAGGAG  120
         ||||||||||||||||||||||||  |||||||||| || ||||||||     || |||||||
Db  1004 TCCTTCACAGCATTGGAAGCTCCCAGAGACAAGTCACTTGACAGTTCACTTGACAGTTCATTGATGAGGAC 1063

Qy   121 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAGAGAGAGAGACTTTCATGCTGCTTCTTTC  180
         |||||||| ||||| | || ||||||  |||||||||||||||||||||||||||||||||
Db  1064 GGAAGTAAATCTGATAGTAGTACCAAACTGCTTAGAGAGAGAGAGACTTTCATGCTGCTTCTTTC 1123

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  240
         || |||| || |||||||||||||| |||||||||||  ||||||||||||||||||||
Db  1124 TGCCGAGGATACATTGCGAGCACAAGGAGGAGATTCAAATGATGTGGAAGAAGTGGAAAGA 1183

Qy   241 CCAGTTCATACTCATGAGGGGGTACATGAGTAGTCGATTGGAACAGGAACATCAGCAGGGTGG  300
         |||||||||| || ||| ||||| ||||| | || ||| | ||||| |||||||||||||||
Db  1184 ACAGTTCATGCTCATGAGGATCATCATGGATCTGACAATCTGACATCTGTCCAGGAACTTGTGG 1243

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
         |||| |||||| ||||||||||||||||| || ||| |||||||||||||||||||||||
Db  1244 TAATGTTCTACAGTTGGGAAGTCAACTAGTTGGAAAAGGAAATTATCAGAAGATGAAGA 1303

Qy   361 AACTGAAGTACAGAGCAGATGAATCTCTAAATTCAAGATGGAATGGCCTCAGGGTAGC  420
         || ||||||||||||| || || ||| ||||||||||||||||||| ||||||||||
Db  1304 AGCTGAAGTGCAAGAACAAATGAATCTCTAAATTCCTAAATTGCCAAGATGCTCAGGGTAGC 1363

Qy   421 TAGGCATGGAAAAACAAAGCAATTACATACAGAGTTTATGGATCTCCAGAATCAGAAACT  480
         |||| |||||||||||||||||||| |||||||||||||||||||||||||||||||| |
Db  1364 TAGGCATGAAAAAACAAAGCAAATTACCAAATTACAAAGTCTAATGGATCTCCAGAATCAGAAATT 1423

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAGAAGGAACTAAGAAAATGGAGGAAGA  540
         |||| || ||||| |||||||||||||| ||| ||||  |||||||| || ||| ||||||
Db  1424 AAAAGAACTAGATGACTGGTTAACAAAAACTGAAGAGAGAAACTGAAGAGAGAACTAAGAAATGGAGGAAGA 1483

Qy   541 GCCCTTCGAGACCTGATCTTGAAGACCAAGACCCAAGAGCCCAAGACGCAGCCAAGTGCTTCA  600
         ||||| ||||||||||| ||||||||||  ||| ||||| | ||||| |||||||||||||||
Db  1484 GCCCTTTGGACCTGATCTTGAAGATCTTGAAGAATGCCAAGACAACAACATAAGGTGCTTCA 1543

Qy   601 AGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCCACTCCACGGTGGTGGTAGT  660
         ||||||||||||| |||| |||||||||||| |||||| ||  ||| ||| ||||||||||||
Db  1544 AGAGATCTAGAACAGGAGACAGTCAGGGTCAGCTCAACTCCTCACTCCACATGGTGGTAGTGGT 1603

Qy   661 TGATGAATCTAGTGGAGATCACGCAACTGCTCGTTGGAAGAACAACTAAGGTATTGGG  720
         ||||||||| ||||  || || ||||||||||||| ||||||||||||||||||||||
Db  1604 TGATGAATCCAGCGGCGTGATCATGCAACAGCGTGCTTGGAAGAACAACTAAGGTACTGGG 1663

Qy   721 AGATCGATGGCCAAACATCTGTAGATGACAGAAGAACCGCTGGCTTCTTTACAAGACAT  780
         |||||||||| |||| |||||| ||||| ||||||||  | |||||| ||||||||||||
Db  1664 AGATCGATGGGCAAATATCTGCAGATGACTGAAGACCGCGAAGAACCGCTGGTTCTTTTACAAGATAT 1723

Qy   781 CCTTCTCAAATGCCAACGTCTTACTGAAGAACCAGTCTTACTGAAGAACTGCTTTTCAGA  840
         ||||  ||| |||| | || ||||| ||||||| ||  || |||||||||||||||||||
Db  1724 TCTTCTAAAATGGCAGCATTTTACTGAAGAACCAGTCTTTTTAGTACATGCGCTTTCAGA 1783

Qy   841 AAAAGAAGAGTGCCAGTGAACAAGATGCACCACACAACTGGCTTTAAAGATCAAAATGAAATGTT   900
```

```
Db   1784  AAAAGAAGATGCAATGAAGACAATTCAGAACAAGTGGCTTAAAGATCAAAATGAAATGAT  1843
            |||||||| |||| ||||  |||| ||||     ||||  ||| ||   |||||||||||
Qy    405  ...TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTC  464

Qy    901  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCCAGAAAAAGAAAAAAGCAATCCAT  960
            | ||||||||||||||    |||| ||||||||||||| |||||||||||||||||||||
Db    181  TGCTTGAGGACACATTGCAAGCAAGCACAAGAGAGAGTTCTAATGATGTGGAAGTGGTGAAAGA  240

Db   1844  GTCAAGTCTTCACAAAATATCTACTTTAAAAATAGATCTAGAAAATGATCCAGAATGCCAACCAT  1903
            ||||||||||| |||||||| |||| |||||||| |||| ||||||||| ||||||||||||
Qy    465  TGCTTGAGGACACATTGCAAGCAAGCACAAGAGAGAGTTCTAATGATGTGGAAGTGGTGAAAGA  524

Qy    961  GGGCAAACTGTATTCACTCAACAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC  1020
            ||||  |||||   ||||| |||||||| ||||| |||| |||||||||||||||
Db    241  CCAGTTCATACTCATGAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG  300

Db   1904  GGAAAACTAAGTTCACTCAATCAATCAAGATCTCATCAGGGACTCGGGCACTGAAAAATAAGTCAGTGAC  1963
            |||  |||| ||||| ||| |||| |||||||||| |||| || |||||||| |||||||||||||||||
Qy    525  CCAGTTCATACTCACTGAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG  584

Qy   1021  CCAGAAGACGGAAGCATGGCTGGATAAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  1080
            |||||||| |||| |||| ||||||| || |||||||||||| ||||| ||| |||||||
Db    301  TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360

Db   1964  TCAAAAGATGGAAATCTGGATGCAAAAACTTGCACAACGTTGGGACAATTTAACCCAAAA  2023
            ||||||||||||  ||||| ||||||| |||||||| |||||||| |||||||  ||||||
Qy    585  TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  644

Qy   1081  ACTTGAAAAGAGTACAGCACAGACT  1105
            ||||||||||||||||| |||||||
Db    361  AACTGAAGTACAAGAGCAGAGCAATGATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420

Db   2024  ACTTGAAAAGAGTTCAGCACCAAATT  2048
            |||||||||||| |||||||| |||
Qy    645  AACTGAAGTACAAGAGCAGAGCAATGATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  704

Qy    421  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  480
            |||||||||||||||||||||||| |||||||||||| ||||| ||||||||||||||
Db    705  TAGCATGGAAAAACAAGCAATTACATAGAGTTTAATGGATCTTCCAGAATCAGAAACT  764

Db    481  GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAGAAGAACAGAGGAAAAATGAGAGAAGA  540
            ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Qy    765  GAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAGAACAGAGGAAAAATGAGAGAAGA  824

Qy    541  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  600
            |||||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||
Db    825  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  884

Db    601  AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGGTGTGGTAGT  660
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Qy    885  AGAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGGTGTGGTAGT  944

Qy    661  TGATGAATCTAGTGGATCACGCAACTGCTCTTGGAAGAACAACTTAAGGTATTGGG  720
            |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    945  TGATGAATCTAGTGGATCACGCAACTGCTCTTGGAAGAACAACTTAAGGTATTGGG  1004

Db    721  AGATCGATGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAAGACAT  780
            ||||||||||||||  ||||| ||||||||||||| |||||||||||||||| |||||
Qy   1005  AGATCGATGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAAGACAT  1064

Qy    781  CCTTCTCAAATGGCAACGTGTTACTGGAAGACAAGTGCCTTTTTAGTGTCATGGTTTCAGA  840
            ||||||||||||||||||||| |||||||||||||||| |||||||||||||||||||
Db   1065  CCTTCTCAAATGGCAACGTGTTACTGGAAGACAAGTGCCTTTTTAGTGTCATGGTTTCAGA  1098

Db    841  AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT  900
            |||| ||||| ||| |||  |||| ||||||| ||  || ||| ||||||||||||||||
Qy   1099  ------------------------------------------------------------  1098

Qy    901  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCCAGAAAAAGAAAAAAGCAATCCAT  960
                                                         |||||||||||||||
Db   1099  ---------------------------------GTTTTAAAAGCGGATCTAGAAAAAGAAAAAAGCAATCCAT  1136

Db    961  GGGCAAACTGTATTCACTCAACAAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC  1020
            ||||||||||| ||||||||| |||||||||||||||||||||||||||||||||||
Qy   1021  GGGCAAACTGTATTCACTCAACAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC  1196

Qy   1021  CCAGAAGACGGAAGCATGGCTGGATAAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  1080
            |||||||| |||| |||| |||||||| || |||||||||||| ||||| ||| |||||
Db   1197  CCAGAGACGGAAGCATGGCCTGGATAACTTTGGCCAGTTCATTGGCAGTTCATTGATGGAGAG  1256

Qy   1081  ACTTGAAAAGAGTACAGCACAGACT  1105
            |||||||||||||| |||||||||
Db   1257  ACTTGAAAAGAGTACAGCACAGATT  1281
```

ESULT 3
Y399453
OCUS   AY399453   5691 bp   DNA   linear   GSS 12-DEC-2003
EFINITION Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
           genomic survey sequence.
CCESSION   AY399453
ERSION   AY399453.1 GI:3975542
EYWORDS   GSS.
OURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE   1 (bases 1 to 5691)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
           Adams,M.D. and Cargill,M.
TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
           gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
EFERNCE   2 (bases 1 to 5691)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
           Adams,M.D. and Cargill,M.
TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
           Rockville, MD 20850, USA
OMMENT   This sequence was made by sequencing genomic exons and ordering
           them based on alignment.
EATURES   Location/Qualifiers
source   1..5691
           /organism="Homo sapiens"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"
gene   <1..>5691
           /locus_tag="HCM0229"
RIGIN

Query Match   40.3%;   Score 846.2; DB 29; Length 5691;
Best Local Similarity 89.1%; Pred. No. 1.7e-177;
Matches 985; Conservative 0; Mismatches 3; Indels 117; Gaps 2;

```
Qy     1  GAGCTATGCCTGCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    294  GAGCTATGCCTACCACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  353

Qy    61  TCCTTCACAGCATTGGAGCTCCTGAAGACAAGTCATTTGGCAGTCATTGGCCAGTGATGAGAG  120
           ||||||||||  |||||||||||                   |||| |||||| |||||||||
Db    354  TCCTTCACAGC------TCCCTGAAGACAAGTCATTTGGCAGTCATTGGCAGTCATTGATGAGAG  404

Qy   121  TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 4
AY399455
LOCUS   AY399455   5697 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
           genomic survey sequence.

```
ACCESSION   AY399455
VERSION     AY399455.1  GI:3755444
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..5697
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>5697
                     /locus_tag="HCM0229"
ORIGIN

Query Match         30.1%;  Score 633.4;  DB 29;  Pred. No. 4e-130;
Best Local Similarity   77.1%;
Matches  852;  Conservative    0;  Mismatches  136;  Indels  117;  Gaps  2;

Qy      1 GAGCTATGGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
          |||||  |  ||||||||||||||||| |||||  |||||||| ||  ||||||||||||
Db    300 GAGTTATGGCTGCTTATGTTGCCACCACCTCCACCAGAGCCCCTA   359

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGCAGTTCATTGATGGAGAG   120
          |||||||||      |||||| ||||||||||||||| ||||||||||||||||||||
Db    360 TCCTTCACAG------GCTCCCAGAGACAAGTCACTTGACAGTTCATTGATGGAGAC   410

Qy    121 TGAAGTAAACCTGGACCGTTATCACAACCTTAGAGAAGTATTATGCTGGCTTCTTTC   180
          |||||||| |||| |  ||||| ||||   |||||||||| |  |  ||||||||||
Db    411 GGAAGTAAATCTGGATAGTTACCAAAACTGCTTTAGAGAAGAAGTACTTTCATGGCTTCTTTC   470

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA   240
          ||| ||||| |||||||| |||||| |||||||||| ||||||| |||| ||||||||
Db    471 TGCCGGAGGATACATTGCGAGCACACAGGAGAGATTTCAAATGATGTTGAAGAGTGAAAGA   530

Qy    241 CCAGTTCATCATGAGGGGTACATGATGAGCCCATCCAGGGGCCGGGTTGG   300
          | ||||||| |||||||  | |||||||| | ||  || ||| ||| ||
Db    531 ACAGTTCATGTCATGAGGGATTCATGATGAGGGATCTGACACGTTGGGTTGTGG   590

Qy    301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAAGAGATGAAGATGAAGA   360
          |||| ||| |||  ||||||  | ||| ||| |||| |||||||||||||
Db    591 TAATGTTCTACAGTTAGGAGAGTCAACTAGTTGAAAAAGGGAAATTATCAGAAGATGAAGA   650

Qy    361 AACTGAAGTACAAGAGACAGATGAATCTCCTAAATTCAAGATGCCTCAGGGTAGC   420
          | |||||||| ||||  ||| ||||||||||||||||||||| |  ||||||||||
Db    651 AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAAGATGGGAATGTCTCAGGGTAGC   710

Qy    421 TAGCATGGAAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT   480
          ||||||||||||| ||| |||| ||||| ||  |||||||||||||||||||||||| |
Db    711 TAGCATGGAAAAACAAGCAAATTACACAAGTTCTAATGATCTCCAGAATCAGAAATT   770

Qy    481 GAAAGAGTTGAATGACTGGCTACAAAAAACAGAAGAACAGAAGAACAAGGAGGAAGA   540
          |||||| ||||| |||||| ||||||||||||| || | | ||||||| |||||||
Db    771 AAAAGAACTAGATGACTGGTTACAAAAACTGAAGACAACCCAAGTACACACATAAGGTGCTTCA   830

Qy    541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACCCAAGTACACACATAAGGTGCTTCA   600
          ||| |||||||||||||||||||| ||||| | ||||||||| ||||||||||||||
Db    831 GCCCTTTGGACCTGATCTTGAAGATCTAAAATGCCAAGTACACAACATAAGGTGCTTCA   890

Qy    601 AGAGGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT   660
          |||  ||||||||||  | |||||| |||||| ||  ||||| ||||| | || |||
Db    891 AGAAGATCTAGAACAGGAGCAGGTCAGGGTCAACTCGCTCACTCACATGGTAGTAGTGGT   950

Qy    661 TGATGAATCTAGTGAGATCACGCAACTGCTGTTGAAGAACAACTAAGGTATTGGG   720
          |||||||||  | | ||||| |||| |||||| ||||||||||| ||||| ||||
Db    951 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTTGGAAGAACAACTTAAGGTACTGGG   1010

Qy    721 AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGTTCTTTTACAAGACAT   780
          ||||||||||||||| ||||| ||||||| | ||||||||| | ||||||||||| ||
Db   1011 AGATCGATGGGCAAATATCTGCAGATGACTGAGACCGCTGGATTGTTTTACAAGATAT   1070

Qy    781 CCTTCTCAAATGCAACGTCTTACTGAAGAAGCAGTGCCTTTTTAGTGCATGGCTTTCAGA   840
          |||||||||||||  | |  ||||||||||| ||
Db   1071 TCTTCTCAAATGCCAGCATTTTACTGAAGAACAG------   1104

Qy    841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT   900
Db   1105 ------   1104

Qy    901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT   960
          ||||  |        ||  |||| |||||    |||| ||||||| |||||||| ||||
Db   1105 ------ACTTTAAAAATAGATCTAGAAAGAAAAAGCCAACCAT   1142

Qy    961 GGGCAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC   1020
          || |||  |  |||||||||| |||||| |||| |||||||| |||  ||||||||||
Db   1143 GGAAAAACTAAGTTCACTCAATCAAGATCTACTTTCGGCACTGAAAAAATAAGTCAGTGAC   1202

Qy   1021 CCAGAAGACGGAAGCGGAGCATGGCTGGATAACTTTGCCCGGTGTTGGGAATAATTTAGTCAAAA   1080
          | |||||| |||| | |||||||||| || |||||||||   | ||||||| |  || || ||||
Db   1203 TCAAAAGATGGAAATCTGGATGGAAAACTTTGCACAACGTTGGGACAATTTAACCCAAAA   1262

Qy   1081 ACTTGAAAAGAGTACAGCACCACAGACT   1105
          ||||||||||||| |||||||  | ||||
Db   1263 ACTTGAAAAGAGTTCAGCACCAAATT   1287


RESULT 5
CB228986
CB228986
LOCUS       CB228986    AGENCOURT 11499247 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
            IMAGE:6884420 5', mRNA sequence.   728 bp   mRNA   linear   EST 10-FEB-2003
DEFINITION  CB228986 AGENCOURT 11499247 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
            IMAGE:6884420 5', mRNA sequence.
ACCESSION   CB228986
VERSION     CB228986.1  GI:28280564
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 728)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Eliot Spindel
            cDNA Library Preparation: CLONTECH
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3135  row: e  column: 11
            High quality sequence stop: 583.
FEATURES             Location/Qualifiers
```

```
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 834)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: The Cepko Laboratory
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM1908  row: e   column: 03
          High quality sequence stop: 796.

FEATURES            Location/Qualifiers
     source          1..834
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5357162"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_94"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."

ORIGIN
Query Match          28.2%;    Score 592.4;  DB 12;   Length 834;
Best Local Similarity 88.4%;   Pred. No. 3.7e-121;
Matches 677;      Conservative  0;   Mismatches  86;   Indels  3;   Gaps  3;

Qy  1065 ATAATTAGTCCAAAACTTGAAAAGAGTACAGCAGATCTCATAGATTACTGCAACAGT 1124
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    44 ATAAAAGTAAGTGAGCAAGAGAGGCTGCTTGGAAGAACTCATAGATTACTGCAGCAGT 103

Qy  1125 TCCCCTGGACCTGAAAGGTTCTTGCCTGGCTTACAGGAGTCTTACAGAAGCTGAAACAACTGCCAATG 1184
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   104 TCCCCTGGACCTGGAGAAGTTTCTTCCTGATTACGGAGCAGCAGAAACAACTGCCAATG 163

Qy  1185 TCCTACAGGATGCTACCCGTAAGGAAAAGGCTCCTAGAAGACTTCCAGGGAGTAAAGAGC 1244
              |||||||||||||||||||||||||||||||||| |||||||||||||||| ||||||
Db   164 TCCTACAGGACGCTTCCCGTAAGGAGGAAGAAGTCCTAGAAGACTCC-AGGGAGTCAGAGAGC 222

Qy  1245 TGATGAACAATGGCAAGAACCTCAAGGTGAAATGAAGCTCCACACAGATGTTATCACA 1304
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   223 TGATGAACAATGGCAAGAATCCCAAGGAGAAATTGAAACTCAAGAGATATCATCACA 282

Qy  1305 ACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCCTGGAAGGTTCCGATGATGCAGTCC 1364
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   283 ATCTTGATGAAAATGGCCAAAAAAATCCTGAGATCCTGGAAGGTTCGGATGATGAAGCAGTCC 342

Qy  1365 TGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGATGAACTTCAGAAAAAAGTCTC 1424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   343 TGTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGATGAACTTCAGAAAAAAGTCCC 402

Qy  1425 TCAACATTAGGTCCCATTGGAAGCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTC 1484
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   403 TCAACATTAGGTCCCATTGGAAGCAGTTCTGACCAGTGGAAGCGTTGCATCTTTCTC 462

Qy  1485 TGCAGGAACTTCGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTA 1544
              |||||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||
Db   463 TTCAGGAACTTCTTGTTGGCTACAGCTGAAAGATGATGAACTGAACTGAGCCGGCACCCA 522

Qy  1545 TTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGCGATGTACATAGGGCCTTCAAGAGGG 1604
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   523 TCGGTGGTGATTTCCCAGCAGTTCAAGAAGCAGAATGATATACATAGGGCCTTCAAGAGGG 582
```

```
source          1..728
                /organism="Macaca mulatta"
                /mol_type="mRNA"
                /db_xref="taxon:9544"
                /clone="IMAGE:6884820"
                /tissue_type="Ovary"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NICHD_Rh_Ov1"
                /note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
                Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.0-4.0 kb. Tissue pooled from
                pre-pubertal, post pubertal sn menopausal monkeys.
                Constructed by Clontech. Note: this is a NICHD Library."

ORIGIN
Query Match          28.4%;    Score 597.2;  DB 14;   Length 728;
Best Local Similarity 97.0%;   Pred. No. 3.1e-122;
Matches 619;      Conservative  0;   Mismatches  18;   Indels  1;   Gaps  1;

Qy   295 GGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGA 354
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGA 60

Qy   355 TGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG 414
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG 120

Qy   415 GGTAGCTAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCA 474
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGTAGCTAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCA 180

Qy   475 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAGGAAAATGA 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GAAACTGAAAGAGTTGAATGACTGGCTGCACAAAAACAGAAGAAGAACAGGAAAATGA 240

Qy   535 GGAAGAGAGCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGT 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAAAGAACCCCTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGT 300

Qy   595 GCTTCAAGAAGATCTAGAACAAGACAAGTCAGGTGAATTCTCTCACTCACATGGTGT 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCTTCAAGAAGATCTAGAACAAGACAAGTCAGGTCAATTCTCTCACTCACATGGTGT 360

Qy   655 GGTAGTTGATGAATCTAGTGAGATCACGCAACTGCTGCTTGGAGAACAACTTAAGGT 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGTAGTTGATGAATCTAGTGAGATCACGCAACTGCTGCTTGGAGAGCAACTTAAGGT 420

Qy   715 ATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTACA 774
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ATTGGGAGATCGATGGGCAAACATCTGCAGATGGACAGAAGACCGCTGGGTTCTTTACA 480

Qy   775 AGACATCCTTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCT 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AGACATCCTTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCT 540

Qy   835 TTCAGAAAAGAGATGCAGTGAACAAGATCACACAACT-GGCTTAAAGATCAAAATG 893
              |||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db   541 TTCAGAAAAGAGATGGGTGGACAAGATCACACAACTCAACTAACTGGGCTTAAGGATCAAAATG 600

Qy   894 AAATGTTATCAGTGCTTCAAAAACTGGCCGTTTTAAAA 931
              ||||||||||||||||||||||||||||||||||||||
Db   601 AAATGTTATCAGTGCTTCAAAAACTGGCCGATGTAA 638
```

```
RESULT 6
BI729851
LOCUS       BI729851    6034951F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
            mRNA sequence.     834 bp    mRNA     linear    EST 20-SEP-2001
DEFINITION  6034951F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
            mRNA sequence.
ACCESSION   BI729851
VERSION     BI729851.1    GI:15706864
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
```

```
Qy   440 AATTTACATAGAGTTTAATGGATCTCCAGAATCAGAATCAGAAACTGAAAGAGAGTTGAATGACTGG  499
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   709 NNTTTACATAGAGTTTAATGGATCTCCAGAATCAGAATCAGAAACTGAAAGAGAGTTGAATGACTGG  768

Qy   500 CTAACAAAAACAGAGAAGAAATGAGGAAGAGCCTCTTGGACCTGATCTT  559
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   769 CTAACAAAAACAGAGAAGAAATGAGGAAGAGCCTCTTGGACCTGATCTT  828

Qy   560 GAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAGGAGTCAATACT  619
         |||||||||||||||||||||||||||||||||||||||||
Db   829 GAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAGGAGTCAATACT  888

Qy   620 CAAGTCAGGTCAATTCTCCACTCACACATGTGGTGGTAGTTGATGAATCTAGTCAGTGGGAT  679
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   889 CAAGTTCAGGTCAATTCTCCACTCACACATGTGGTGGTAGTTGATGAANNNANTGGAGAC  948

Qy   680 CACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATC  739
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   949 CACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATC  1008

Qy   740 TGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGT  799
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1009 TGTAGGTGGACAGAAGCCCGCTGGGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGT  1068

Qy   800 CTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGAAAAAGAGAGATGCAGTGAAC  859
         |||||||||
Db  1069 CTTACTGAAGAACAG------------------------------------------  1083

Qy   860 AAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTTATTCAAGTCTTCAAAAACTG  919
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1084 ------------------------------------------------------------  1083

Qy   920 GCCGTTTAAAAGGGGATCTAGAAAGAAAAAAGCAATCCATGGGCAAACTGTATTCACTC  979
         |||||||||||||||||||||
Db  1084 ---GTTTAAAAGGGGATNNNGAAAGAAAAAAGCAATCCATGGGCAAACTATATTNNNNN  1140

Qy   980 AAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGACCCAGAGACGGAAGCATGG  1039
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 NNACAAGNTNTTCNTNTNAACACTGANGAATAAGTCAGTGACCCNNAAGACGGAAGCNNGG  1200

Qy  1040 CTGGATAACTTGCCCGGTTGTTGGAATAATTTAGTCCAAAAACTTGAAAAAGTACAGCA  1099
         |||||||||||||||||||||||||
Db  1201 NNGGATAACTTGCCCGNNNTTGGATAATTTNNNCCAAAAACNNGAAAAAGAGTACAGCA  1260

Qy  1100 CAGACT  1105
         ||||||
Db  1261 CAGATT  1266
```

RESULT 8
CB177816/c
LOCUS   CB177816   595 bp   mRNA   linear   EST 31-JAN-2003
DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
            similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ;, mRNA sequence.
ACCESSION   CB177816
VERSION   CB177816.1 GI:28186206
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 595)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE   Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT   Other ESTs: is21c01.y1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

```
Qy  1605 AATTGAAAAACTAAAGAACCTGTAATCATAGAGTACTCTTGAGACTGTACGAATAATTCTGA  1664
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   583 AATTGAAAAACTAAAGAACCTGTAATCATAGAGTACTCTTGAGACTGTACGAATAATTCTGA  642

Qy  1665 CAGAGCAGCCTTGGAAGGACTAGAGAAACTC-TACCAGGAGCCCAGAGAGCTGCCTCCT  1723
         ||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db   643 CAGAGCAGCCTTGGAAGGACTAGAGAAACTCTTACCAGGAGCCCAGAGAACTGCCTCCT  702

Qy  1724 GAGGAGAGAGCCCAGAAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACT  1783
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   703 GAAGAAAGAGCTCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAAGAGGTCAACGCT  762

Qy  1784 GAGTGGGAAAAATGAACCTGCACTC-CGCTGACTGGCGCAGAGAAA  1828
         |||||||||||||||||||||||||  ||||||||||||||||||||
Db   763 GAATGGGAAAAATGAACCTGCGCTGCAAGCTGATTGGCAGAGAGAAA  808
```

RESULT 7
Y399454
LOCUS   AY399454   5676 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY399454
VERSION   AY399454.1 GI:39755443
KEYWORDS   GSS.
SOURCE   Pan troglodytes (chimpanzee)
  ORGANISM Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 5676)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 5676)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..5676
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene   <1..>5676
                     /locus_tag="HCM0229"
ORIGIN

```
Query Match         24.5%;  Score 514.4;  DB 29;  Length 5676;
Best Local Similarity  75.2%;  Pred No. 1.3e-103;
Matches  636;  Mismatches  0;  Indels 108;  Gaps  1;

Qy   260 GGGTACATGATGGATTTGACAGCCCCATCAGGGCCGGGCGGCCTGGTAATATTCTACAATTGGGA  319
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   529 GGGTACATGATGGATNNNNNAGCCCCATCAGGGCCGGGCGGTGGTAATANNNNANTTGGGA  588

Qy   320 AGTAAGCTGATTGGAACACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAG  379
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   589 AGTNNNNTGATTGGAACAGCAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAG  648

Qy   380 ATGAATCTCCTAAATTCAAGATGGGAAATGGGAANNNNNNNNNNNNNNNNNNNNNN  439
         ||||||||||||||||||||||||||||||||||
Db   649 ATGAATCTNNNNNATTCAAGATGGGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNN  708
```

RESULT 9

```
LOCUS       CB850319                1047 bp    mRNA    linear   EST 01-SEP-2003
DEFINITION  MRA-0070 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA sequence.
ACCESSION   CB850319
VERSION     CB850319.1  GI:34380806
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1047)
  AUTHORS   Yu,J., Farjo,R., MacNee,S.P., Baehr,W., Stambolian,D.E. and
            Swaroop,A.
  TITLE     Annotation and analysis of 10,000 expressed sequence tags from
            developing mouse eye and adult retina
  JOURNAL   Genome Biol. 4 (10), R65 (2003)
  MEDLINE   22881944
   PUBMED   14519200
COMMENT     Contact: Swaroop, A.
            Department of Ophthalmology and Visual Sciences
            Kellogg Eye Center, University of Michigan
            540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
            Tel: 734 615 2246
            Fax: 734 647 0228
            Email: swaroop@umich.edu.
FEATURES             Location/Qualifiers
     source          1..1047
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="retina"
                     /clone_lib="MOUSE ADULT RETINA"
                     /note="Vector: pSPORT1"
ORIGIN
```

Query Match　　　23.5%;　　Score 493.2;　DB 14;　Length 1047;
Best Local Similarity　83.2%;　　Pred. No. 4.8e-99;
Matches　610;　Conservative　0;　Mismatches　118;　Indels　5;　Gaps　5;

```
Qy   133 GGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTCTGCTGGGAGGACAC 192
         ||||||||||||||||||||||||| |||||||| |||||||||||||||| || |||||
Db    86 GGATAGTTACCAAACTGCTTTANAAGAAGTACTTTCATGGCTTCTTCTGCCAGGGATAC 145

Qy   193 ATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTTCATAC 252
         ||||||||||||||||||||||||||||||||  ||||||||||||||||| |||||||||
Db   146 ATTGCGAGCACAAGGAGAGATTTCAAATGATGTGTGAAGAAGTGAAAGAACAGTTTCATGC 205

Qy   253 TCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGGTAATATTCTACA 312
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206 TCATGAGGGATTCATGATGATCTGACATCTCATCAAGGACTTGTTGGTAATGTTCTACA 265

Qy   313 ATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGAAAACTGAAGTACA 372
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266 GTTAGGAAGTCAACTAGTTGGAAAAGGAAATTATCAGAAGATGAAGAAGAGCTGAAGTGCA 325

Qy   373 AGAGCAGGATGAAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAG-CTAGCAGGAAA 431
         |||| ||| ||||||||||||||||||||||||||||| || ||||||| ||| |||||||
Db   326 AGAACAAATGAATCTCCTAAATTCAAGATGGGAATGTCTCAGGGTANCCTAGCATGGGAAA 385

Qy   432 AACAAAGCAATTACATAGAGTTTAATGAGTTTCAATGATCTCCAGAATCAGAAAGAGTTGA 491
         |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db   386 AACAAAGCAATTACCAACAAGTTCAATGATCTCCAGAATCAGAAATTAAAAGAACTAG 445

Qy   492 ATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGAGCCTCTTGGAC 551
         ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   446 ATGACTGGTTAACAAAAACTGAAGAAGAGACTCAANAAAAATGGAGGAAGAGCCCTTTGGAC 505

Qy   552 CTGATCTTGAAGACCTAAACGCCAAGTACAACACATAAGGTGC-TTCAAGAAGATCTA 610
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   506 CTGATCTTGAAGACTTCAAAAATGCCAAGTACAACACATAAGGTGCTTTCAAGAAGATCTA 565
```

```
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.
FEATURES             Location/Qualifiers
     source          1..595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6553129"
                     /tissue_type="purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN
```

Query Match　　　24.2%;　　Score 509;　DB 14;　Length 595;
Best Local Similarity　100.0%;　Pred. No. 1.3e-102;
Matches　509;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

```
Qy     1 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTGACCCTACACGGAGCCCATT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   509 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 450

Qy    61 TCCTTCACAGCAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   449 TCCTTCACAGCAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 390

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   389 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC 330

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   329 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 270

Qy   241 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   269 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 210

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   209 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 150

Qy   361 AACTGAAGTACAAGAGCAGGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   149 AACTGAAGTACAAGAGCAGGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 90

Qy   421 TAGCATGGAAAAAACAAAGCAATTACATAGAGTTTAATGAGTTTCAATGATCTCCAGAAACT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    89 TAGCATGGAAAAAACAAAGCAATTACATAGAGTTTAATGAGTTTCAATGATCTCCAGAAACT 30

Qy   481 GAAAAGAGTTGAATGACTGGCTAACAAAAA 509
         ||||||||||||||||||||||||||||||
Db    29 GAAAAGAGTTGAATGACTGGCTAACAAAAA 1
```

```
Qy   611 GAACAAGAACAAGTCAGGTGTCAATTCTCTCACTCACACATGGTGGTGAGTTGATGAATCT 670
         ||||||   || |||||||||||||||| | ||||||||||||| | |  |||||||||||
Db   566 GAACAGGAGGAGCAGGTCAGGTGTCAACTCGCTCACTCACACATGGNAGTAGTGGGTGATGAATCC 625

Qy   671 AGTGGAGATCACCGGAACTGCTGCTTTGGAAGAACAAC-TTAAGGTATTGGGAGATCGATG 729
         || | |||||| || ||||||||||||| |||| ||||| ||||||| ||||||||||||||
Db   626 AGCGGGGATCATCGAACAGCTGCTTTGGAAANAACAACTTTAAGGNACTGGGAGATCGATG 685

Qy   730 GGCAAACATCTG-TAGATGGACAGAAGACCGCTGGGT-CTTTTACAGAGACATCCTCTC 787
         |||||| ||||| | ||| |||| ||| ||||  ||| |||||||||  ||| |||||||
Db   686 GGCAAATATCTGCGAAAATGGACTGAAAACCGTTGGATTGTTTTTACANGATTTTCTTCTA 745

Qy   788 AAATGGCAACGTCTTACTGAAGAAACAGTGCCTTTTAGTGCATGGCGTTTCAGAAAAAGAA 847
         |||||||| |   ||||||||  ||  | |||||||||  ||  ||   ||  |||||||||
Db   746 AAATGGCAGCNTTTTACTGAAAAANANNGCCTTTTATNCANGGNTTTTAAAAAAAAAAA 805

Qy   848 GATGCAGTGAACA 860
         | || || | ||
Db   806 TGCATTGANAACA 818
```

```
RESULT 10
BI730168
LOCUS       BI730168      750 bp    mRNA     linear    EST 20-SEP-2001
DEFINITION  603349711F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357187 5',
            mRNA sequence.
ACCESSION   BI730168
VERSION     BI730168.1  GI:15707181
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 750)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11908 row: f column: 04
            High quality sequence stop: 747.
FEATURES             Location/Qualifiers
     source          1..750
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5357187"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_94"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN
```

```
Query Match   23.4%;  Score 491.6;  DB 12;  Length 750;
Best Local Similarity 89.3%;  Pred. No. 1e-98;
Matches 598;  Conservative 0;  Mismatches 64;  Indels 8;  Gaps 6;

Qy  1098 CACAGACTCATAGATTACTGCAACAGTTCCCCTGACCTGAAAAGTTCTTGCCTGG- 1156
         |  ||| ||||||||||||||||| |||||||  | |||||| ||||| ||||| ||
Db    80 CAGAGAACTCATAGATTACTGCACAGTTCCCTGGACCTGGACAAGTTTCTTTCCTGGA 139

Qy  1157 CTTACAGAAGCTGAAAC--AACTGCCAATGCTCCTACAGGATGCTACCCG-TAAGGAAAGG 1213
```

```
Db   140 TTTACGGACAGAAACAGAGTCCCAATGTCCTACAGGACGCTTCCCGCTAAGGAGAAG 199
Qy  1214 CTCCTAGAAGACTCCAAGG--AGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAG 1271
Db   200 CTCCTAGCAGACTCCAAGGGCAGTCAGAGAGTGATGAAACCATGGCAAGATCTCCAAG 259
Qy  1272 GTGAAA-TTGAAGCTCACACAGATGTTATCACAACTGGATGAAAACAGCCAAAAATC 1330
Db   260 GAGAAAGTTGAACTCACACAGATCTCATCACAATCTTGATGAAAATGGCCAAAAATC 319
Qy  1331 CTGAGATCCCTGGAAGGTTCCGATGATGCAGTCGTCCTGTTTACAAAGACGTTTGGATAACATG 1390
Db   320 CTGAGATCCCTGGAAGGTTCGGATGATGAAGCACCCCTGTTTACAAAGACGTTTGGATAACATG 379
Qy  1391 AACTTCAAGTGGAGTGAACTGTCTCCAACATTAGGTCCCATTTGGAAGCC 1450
Db   380 AAATTCAAGTGGAGTGAACTTCAGAAAAAGTCTCCAACATTAGGTCCCATTTGGAAGCA 439
Qy  1451 AGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTTCTGGTGTGGCTACAG 1510
Db   440 AGTTCTGACCAGTGGAAGCGTTTGCATCTTCTCTGCAGGAACTTCTTCTGTGTGGCTACAG 499
Qy  1511 CTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAG 1570
Db   500 CTGAAAGATGATGAACTAGGCCGTCAGCCATCCGTGGTGATTTCCCAGCAGTTCAG 559
Qy  1571 AAGCA-GAACGATGTACATAGGGCCTTCAAGAGGATTGAAAACTAAAGAACCTGTAAT 1629
Db   560 AAGCAGCAATGATGATATACATAGGGCCTTCAAGAGGATTGAAAACTAAAGAACCTGTAAT 619
Qy  1630 CATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCCTTTGGAAGGACTAGA 1689
Db   620 CATGAGTACTCTTGGGACTGTGGAAATATTTCTGACAGAGCGCCTTTGGAAGGACTAGA 679
Qy  1690 GAAACTTCACCAGGAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGCCCAGAATGTCACTCG 1749
Db   680 GAAACTTCACCAGGAGGAGCCCAGAGAGCTGCCTCCTGAGAAGAGAGCTCAGAATGTCACTCG 739
Qy  1750 GCTTCTACGA 1759
Db   740 GCTCCTACGA 749
```

```
RESULT 11
AL121550
LOCUS       AL121550      579 bp    mRNA     linear    EST 04-SEP-2003
DEFINITION  DKFZp762L078_r1 762 (synonym: hmel2) Homo sapiens cDNA clone
            DKFZp762L078 5', mRNA sequence.
ACCESSION   AL121550
VERSION     AL121550.1  GI:5927551
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 579)
AUTHORS     Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
TITLE       EST (Ottenwaelder, et al.)
JOURNAL     Unpublished (1999)
COMMENT     Contact: MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by MediGenomix (Martinsried/Germany) within the cDNA
            sequencing consortium of the German Genome Project. No 81 sequence
            available.
            This clone (DKFZp762L078) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
```

FEATURES
    source   Location/Qualifiers
                    1..579
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="DKFZp762L078"
                    /tissue_type="melanoma (MeWo cell line)"
                    /dev_stage="adult"
                    /lab_host="DH10B"
                    /clone_lib="762 (synonym: hmel2)"
                    /note="Vector: pSportI; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match   23.2%;   Score 486.6; DB 9; Length 579;
Best Local Similarity 97.2%;   Pred. No. 1.3e-97;
Matches 495;   Conservative   0;   Mismatches 14;   Indels   0;   Gaps   0;

Qy   1260 AAGGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACA 1319
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 AGGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACA 130

Qy   1320 GCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTT 1379
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 GCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGAGACGTT 190

Qy   1380 TGGATAACATGAACTTCAAGTGGAGTGAGTGAACTTCCGAAAAAGTCTCTCAACATTAGGTCCC 1439
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 TGGATAACATGAACTTCAAGTGGAGTGAGTGAACTTCCGAAAAAGTCTCTCAACATTAGGTCCC 250

Qy   1440 ATTTGGAAGCCAGTTCTGACCAGTGGAAGGGTCTGCACCTTTCTGCAGGAACTTCTGG 1499
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 ATTTGGAAGCCAGTTCTGACCAGTGGAAGGGTCTGCACCTTTCTGCAGGAACTTCTGG 310

Qy   1500 TGTGGCTACAGCTGAAAGATGAAGATGAATTAAGCCCGGCAGGCACCTATTGGAGGCGGACTTTC 1559
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 TGTGGCTACAGCTGAAAGATGAAGATGAATTAAGCCCGGCAGGCACCTATTGGAGGCGGACTTTC 370

Qy   1560 CAGCAGTTCAGAAGCAGAAGCAGAAGAAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAG 1619
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 CAGCAGTTCAGAAGCAGAAGCAGAAGAAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAG 430

Qy   1620 AACCTGTAATCATGAGTACTCTTGAGACTGTACGAAGATATTTTCTGACAGAGCAGCCTTTGG 1679
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 AACCTGTAATCATGAGTACTCTTGAGACTGTACGAAGATATTTTCTGACAGAGCAGCCTTTGG 490

Qy   1680 AAGGACTAGAGAAAACTCTACCAGGAGGAGCCCAGAGAGCTGCCTCCTGAGAGGAGAGCCCAGA 1739
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 AAGGGCTAGAGAAAACTCTACCAGGAGGAGCCCAGAGAGCTGCCTCCTGAGAGGAGAGCCCCAG 550

Qy   1740 ATGTCACTCGGCTTCTACGAAAGCAGGCT 1768
                 ||||||||||||||||||||||
Db   551 AATGTCACTCGGCTTCTACGAAAGCAGGCT 579

ORIGIN

Query Match   20.2%;   Score 424.2; DB 12; Length 663;
Best Local Similarity 78.1%;   Pred. No. 1e-83;
Matches 518;   Conservative   0;   Mismatches 144;   Indels   0;   Gaps   1;

Qy   1243 GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCA 1302
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCTCATGAAGCAGTGCAGTGCAGGATCTACAGGCAGAGAAATTGATGCACATACTGACATCT-TNN 59

Qy   1303 CAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 1362
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   60 CAACCTGGATGAAAACGGGCAGAAAATCCCTGAGATCCCTGGAAGGGCTCAGAGGATGTCGT 119

Qy   1363 CCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTC 1422
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120 CCTGTTGCAGAGACGTCTCGATAACATGAACTTCAGATGGAGTGAGCTTAGGAAGAAATC 179

Qy   1423 TCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTC 1482
                 ||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 TCTAAACATTAGATCTCATTTGGAAGCCAGTTCTGACGACCAGTGGAAGCGTTTACATCTCTC 239

Qy   1483 TCTGCAGGAACTTCTGTGTGGCTGCTGCAGGCTGCAATTGAGAAGCAGAACATGGGGCTTCAAGAG 1542
                 |||||||||||||||||||||||||||||||||||||||||||||||
Db   240 TCTTCAGGAACTTTTGGCATGGCCTGCTGCAATTGAAGAAGGAGATGAATTAAAACAGCAAGCACC 299

Qy   1543 TATTGGAGGCGGACTTTCAGCAGTTCAGAGAAGCAGAACATGTACATAGGGCCTTCAAGAG 1602
                 ||||||||||||||||||||||||||||||||||||||||||||||||
Db   300 CATTGTGGAGATATTCCCACTGTGCAGAGACAGAATGATGTTCATAGGACTTTCAAGAG 359

Qy   1603 GGAATTGAAAACTAAAGACCTGTAATCATGAGACCTGTTATCATGAATGCACTTGACTGTGCGCACTCTTCCT 1662
                 ||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 GGAGCTGAAACAAACAAAGACCTGTTATCATGAATGACCTGTGTGCGACTGTCTTCCT 419

Qy   1663 GACAGAGCAGCCTTTGGAAGGACGTTTGGAAGGACTAGAGAAACTTCACCAGGAGGCCCAGAGAGCTGCCTCC 1722
                 ||||||||||||||||||||||||||||||||||||||||||||||||
Db   420 GGCAGATCAACCAGTAGAGGGAGCTGGAAAAGGTCTATCCGAGAAACAGCAGGATTATCACC 479

Qy   1723 TGAGGAGAGAGCCCAGAAGTGCACTCGGCTTCACGAAAGCAGGCTGAGGAGGTCAATAC 1782
                 |||||||||||||||||||||||||||||||||||||||||||||||
Db   480 TGAGAGAGGGCCCAGAAGTGCACTCAAAGTTCTCCGAGGCAAGCAGCAGATGATGTCAGNAC 539

Qy   1783 TGAGTGGGAAAATTGAACCTGCACTCCCGTACTCCCGTACTCCGACTGGCGAGAGAAAAATGAGAGACCCT 1842

RESULT 12
BM488464

LOCUS   BM488464   663 bp   mRNA   linear   EST 07-FEB-2002
DEFINITION pgm2n.pk007.117 Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
            clone pgm2n.pk007.117 5' similar to sp|P11533|DMD_CHICK Dystrophin
            pir|S02041 dystrophin, muscle - chicken emb|CAA31746.1| (X13369)
            dystrophin (AA 1 - 3660) [Gallus gallus], mRNA sequence.
ACCESSION   BM488464
VERSION   BM488464.1   GI:18609395
KEYWORDS   EST.
SOURCE   Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 663)
AUTHORS   Cogburn,L.A. and Monsonego-Ornan,E.
TITLE   ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and

FEATURES
    source   Location/Qualifiers
                    1..663
                    /organism="Gallus gallus"
                    /mol_type="mRNA"
                    /strain="Commercial broiler and Ottawa Res. Centre
                    strains 90 & 21"
                    /db_xref="taxon:9031"
                    /clone="pgm2n.pk007.117"
                    /sex="Male and Female"
                    /tissue_type="Breast muscle, leg muscle and epiphyseal
                    growth plate"
                    /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
                    ,11 weeks);growth plate(1d,7d,14d post-hatch)
                    /lab_host="E. coli EMDH10B"
                    /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
                    and Epiphyseal Growth Plate cDNA library (pgm2n)"
                    /note="Vector: pCMVSPORT6; Library made from equivalent
                    pools of total RNA isolated from each tissue (embryonic
                    muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
                    plate 33.3% of the final RNA pool). Single pass sequencing
                    from 5'-end"

```
Db      7 CGAGCACGACCAGTGGAAGCGTTTACATCTCTCTTCAGGACTTTGGCATGGCTGC 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1509 AGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTC 1568
              | |||||| ||||  | |||  || |   || ||   || ||||||||| |||||||
Db     67 AATTGAAGGAGGATGAATTAAACAGCAGGCACCCATTGGTGAGATATTCCCACTGTGC 126

Qy   1569 AGAAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAAAACTAAAGAACCTGTAA 1628
              |||||||||||| | ||  |||||  ||||||| || || ||  || ||||||| ||
Db    127 AGAAGCAGAATGATGTGTTCATAGGACTTTCAAGAGGAGAGCTGAAAACAAAAGAACCTGTTA 186

Qy   1629 TCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAG 1688
              ||||| |||||| |||| | || ||  | || |||  ||||| ||||| |||||||
Db    187 TCATGAATGACTTGAGACTGTGCGCGACTCTTCCTGGCAGATCAACCAGTAGAGGGACTCG 246

Qy   1689 AGAAACTTACCAGGAGCCCAGAGACTGCCTCCTGAGAGAGCAGAGAAATGTCACTC 1748
              |||| ||| ||||||  ||   ||  | |||   | |||||||  ||    ||||
Db    247 AAAAGCTCATCCAGAACCAAGACCTATCACCTGAGAGAGAGCGCCCAGAATGTCACTA 306

Qy   1749 GGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACT 1808
              |||||||||| | || || ||| || ||  | ||||| |||| | ||||||| ||
Db    307 AAGTTCTCCGAGGCAGCAAGGTCAGCAGATGATGTCAGAACTGAGTGGGATAAGCTAAATCTACGTT 366

Qy   1809 CCGGTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTCAAGAGG 1868
              | | |||| || ||||  |  |||||||| |  ||||||| |||||| | |||||||
Db    367 CTGCTGATTGGCAAAAGAAGAGATAGGATGATGTCTCTTGAAAGACTGCAGGGTCTTCAGGAGG 426

Qy   1869 CCACGGATGAGCTGAACCTCAAGCTGCGGCCAAGCTGAGTGATCAAGGGATCCTGGCAGC 1928
              | | |||| || || || ||||| |  ||||| | || || || |||||| || ||
Db    427 CAATGATGAACTAGACCGAGAACTGCGCGGCCCAGGTGAAGCATTCAAGGGATCCTGGCAGC 486

Qy   1929 CCGTGGGCGGATCTCCTCATTGACTTCTCTCCAGGATTCACTCGAGAAAGTCAAGGCACTTC 1988
              | | ||| ||||||||  |||| ||| || ||||| |||||| | |||| ||||| ||||
Db    487 CAGTGGGGGGATCTGCTGATAGACTTCTCTCCAGGATCACTTAGAGAAAAAGTCAAGGTTTATC 546

Qy   1989 GAGGAGAAATTGCGCCTGCAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGC 2048
              |||| |||||  |  ||||  | ||||| ||| | || | || ||  || ||| |||
Db    547 GAGCAGAAATTGGTGCCCCTTTANAGAGAAGGTGCATCAAGTCAATGAGTCTGGGCTCACCCGT 606

Qy   2049 TTACACTTTGGGCATTCAGCTCTCCACCGTATAACCTC 2086
              ||||||| || |||||| || || |||   |||   |
Db    607 TCGCTCCCCCTGATATTCAGTTCTCCCATACACTCTC 644
```

```
Db    540 TGAGTGGGATAAGCTAAATCTACGTTCTGCTGATTGGCAAAAGAAGAGATGATGCTCT 599
              |||||||| |  | ||| | | |  ||  ||  |||  |  |||||| ||||| |
Qy   1843 TGAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGC 1902
              ||||||||||  |||||  |  ||| || || ||||| || || ||||||| ||
Db    600 TGAAAGACTTGCCAGGGTCTTCAGGAGGCAATGATGAACTGAGGAGGCNNNCCTGAAACTGCGCGCCAGGC 659

Qy   1903 TGA 1905
              |||
Db    660 TGA 662
```

RESULT 13
U313510
LOCUS   BU313510   644 bp   mRNA   linear   EST 28-NOV-2002
DEFINITION 603540290F1 CSEQCHN61 Gallus gallus cDNA clone ChEST508c24 5', mRNA
            sequence.
ACCESSION   BU313510
VERSION   BU313510.1 GI:25821511
KEYWORDS   EST.
SOURCE   Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 644)
AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE   A Comprehensive Collection of Chicken cDNAs
JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE   22335534
PUBMED   12445392
COMMENT   Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES   Location/Qualifiers
     source   1..644
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Compton Line 15I"
                     /db_xref="taxon:9031"
                     /clone="ChEST508c24"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN61"
                     /note="Organ: heart; Vector: pBluescript II KS(+); Site_1:
                     EcoRI; Site_2: NotI; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."

ORIGIN

```
Query Match       18.5%;   Score 389;   DB 13;   Length 644;
Best Local Similarity  75.5%;   Pred. No. 7e-76;
Matches 482;   Conservative  0;  Mismatches 156;  Indels  0;  Gaps  0;

Qy   1449 CCAGTTCTGACCAGTGGAAGCGTCGCACCTTCTCTGCAGGAACTTCTGGTGTGGCTAC 1508
              | || ||||||| |||  |||  |||| ||| |  |||||| | | ||| |||| ||
```

RESULT 14
BC011062
LOCUS   BC011062   2334 bp   mRNA   linear   HTC 04-MAR-2003
DEFINITION Mus musculus, Similar to utrophin, clone IMAGE:3979320, mRNA.
ACCESSION   BC011062
VERSION   BC011062.1 GI:15029695
KEYWORDS   HTC.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 2334)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (25-JUL-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 23 Row: j Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein
This clone has the following problem: retained intron.

```
FEATURES             Location/Qualifiers
     source          1..2334
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3979320"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match       16.3%;  Score 343;  DB 11;  Length 2334;
Best Local Similarity   59.0%;  Pred. No. 1.6e-65;
Matches    589;  Conservative   0;  Mismatches  410;  Indels   0;  Gaps   0;

Qy   104 AGTTCATTGATGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTA 163
         ||||||||  |||||| | ||| ||||||  ||| ||||||||| |||||| ||||
Db  1042 AGCACGTCACTGAGTGACATGATTGGACAGCAGCAGCGTAGCGCTAGCAGCGAAGTG 1101

Qy   164 TTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGAT 223
Db  1102 CTGACGTGGCTGCTGTCCGCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 1161

Qy   224 GTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGGTACATGATGATTGACAGCC 283
Db  1162 GTCGAAGAGTCAAAGAGCAGTTTGCTACCCATGAAACTTTTATGATGGGAGCTGACAGCA 1221

Qy   284 CATCAGGGGCCGGGTTGGTAATATTCTACAAATTGGAAGTAAGCTGATTGGAACAGGAAAA 343
Db  1222 CACCAGCAGCAGCGTGGGAGCGTCCTGCAGGCTGGCAGCAGGATGAATCTCCTAAATTCAAGATGG 1281

Qy   344 TTATCAGAAGATGAAGAAAACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGATGG 403
Db  1282 CTGTCAGAGGAGGAGGAGTTGAGATCCAGGAACCAGGAACAGATGACCTTGCTGAATGGCAAGGTGG 1341

Qy   404 GAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATGGAT 463
Db  1342 GAGGGCGCTCCGGGTGAGAGCATGGAGAGGCAGTCCCGGCTGCACGCGTCTCTGATGGAG 1401

Qy   464 CTCCAGAATCAGAAACTGAAAGAGTTGAAATGACTGGCTAACAAAAACAGAGAAGAAGAACA 523
Db  1402 CTGCAGAAGAAAACAGCTGCAGCTCTCAGCTCTTCAAGCTGGTGGCTGGCCCCTCCAAGCTCTGGCCCATT 1461

Qy   524 AGGAAAATGGAGGAAGAGAGCCTCTGGAACCTGGACCTGATCTTGAAGACTAAAACGCCAAGTACAA 583
Db  1462 CAGAGATGATGGATTGTGGATGAAAAACAGTGGGGAGAGTGCCACAGTGCTCCCAGCTCCTCTCTGCAGGGTAAAT 1521

Qy   584 CAACATAAGTGCTTCAGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTCAATTCTCACT 643
Db  1522 GAACATAAAGTTTGCAAAATGACCTTGAAGCTGAAGGTGAAGGTAAATTCCTTAACT 1581

Qy   644 CACATGGTGGTGGTAGTTGATGAATCTAGTGGAATCTAGTGGAGATCACGCAACTGCTCTTTGGAAGAA 703
Db  1582 CACATGGTCGGTGATTGTGATGAAAAACAGTGGGGAGAGTGCCACAGTGCTCTTCGTCTCTGCAGGAAGAT 1641

Qy   704 CAACTTAAGTGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGG 763
Db  1642 CAGTTACAGAACTGGTGAGCGCTGGACAGCTGTAATGCCGCGTGATCGTCTTGGACTGAAGAACGTTGG 1701

Qy   764 GTTCTTTTTACAAGACATCCTTCTCAAATGCAACGTCCTTACTGAAGAACAGTGCCTTTTT 823
```

```
Qy   824 AGTGCATGCGTTTCAGAAAAAGAAGATGCAGTGAACAGATTCACCAACTGGCTTAAA 883
         |||||  ||||||  ||||||||||||||||||  |  |  |||||||||| ||||
Db  1762 GAGGCTTGGCTACCGAAAAGGAAGAGGCTTGAATAAAGTTCAAACAGCAACTTTAAA 1821

Qy   884 GATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGCGCGTTTAAAAGCCGATCTAGAA 943
Db  1822 GACCAGAGGAGAACTAAGTGCAGTGTCCAGTGTCCGCGCTCCGCTATATTGAAGAGACATGGAA 1881

Qy   944 AAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAACAAGAGATCTTCTTTCAACACTG 1003
Db  1882 ATGAAGAGGCAGACTCTGGATCAACTGAGTGACATTGGCCAGGATGTGGGCCAATTACTC 1941

Qy  1004 AAGAATAAGTCAGTGACCCAGAAGACGGAAGCAGTGCTGGATAAACTTGCCGGTGTTGG 1063
Db  1942 AGTATTCCCAAGGCATCTAAGAACAGTGACCTCTGAGGAGCTAACACAGAGACATGGAA 2001

Qy  1064 GATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACAG 1102
Db  2002 GATTCTCTGGTTCAGAGACTCTCGAAGACTCTTCTAACCAG 2040
```

RESULT 15
AY407022

```
LOCUS       AY407022                9915 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Homo sapiens UTRN gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY407022
VERSION     AY407022.1  GI:39762993
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9915)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671102
REFERENCE   2  (bases 1 to 9915)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..9915
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>9915
                     /gene="UTRN"
                     /locus_tag="HCM2752"
ORIGIN

Query Match       15.9%;  Score 334.6;  DB 29;  Length 9915;
Best Local Similarity   62.5%;  Pred. No. 1.5e-63;
Matches    523;  Conservative   0;  Mismatches  314;  Indels   0;  Gaps   0;

Qy   103 CAGTTCATTGATGAGGAGTGGAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGT 162
         ||||  | |||| |||||| | | ||||  ||  ||||||||||||||||||||||
Db   918 CAGCACTGTCACTGAGGTTGACATGGATCTGACACTGGCTATCAGATTGCGTTGGAGGAAGT 977
```

```
y    163 ATTATCGTGCTTCTTCTCGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGA  222
            ||||| |||||| ||| | |||||||||||  | |||  ||||||  |||||| ||||||
b    978 GCTGACCTGTTGCTTTCTTGCTGCTGAGGACACCTTTCCAGGAGCAGGATGATATTTCTGATGA 1037

y    223 TGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGGATTTGACAGC  282
            | |||||||||||  ||||||||||| |||| |||| ||||||| ||||| |||||||||
b   1038 TGTTGAAGAAGTCAAAGACCCAGTTGCAACCCATGAAGCTTTTATGATGGAACTGACTGC 1097

y    283 CCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAA  342
            || || |||||| ||| ||||| ||| | ||||||| |||||  ||| ||||| ||||||
b   1098 ACACCAGCAGCAGTGTGGGCAGCGTCCTGCAGCGCAGCCAACCACTGATAACACCAAGGAAC 1157

y    343 ATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATG  402
            | |||||||||||| | |||| ||| |||||||||| |||||||| |||||||||||||
b   1158 TCTGTCAGACGAAGAAGATTTGAGATTCAGGAACAGATGACCCTGCTGAATGCTAGATG 1217

y    403 GGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGA  462
            || || |||| ||||| |||| ||||||||| |||| | ||| |||||||| ||||||||
b   1218 GGAGGCTCTTAGGGTGGGACAGTATGGACAGTCCCGGCTGCACGATGTGCTGATGGA 1277

y    463 TCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAAC  522
            |||| ||||| || |||| |  || |||||| ||||| ||||||||| ||||| ||||||
b   1278 ACTGCAGAAGAAGCAACTGCAGCAGCTCTCCGCCTGGTTAACACTCACAGAGGAGCGCAT 1337

y    523 AAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACA  582
            || |||||||||  |||   ||| |||| |||||| ||||||||| ||| ||||||||
b   1338 TCAGAAGATGGAAACTGCCCCTGGATGATGATGAAAATCCTACAAAAGCTGCTAGA 1397

x    583 ACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCAC  642
            || || |||||| ||||| || || || |||| |||||||||| |||| |||| |||||
b   1398 AGAACATAAAAGTTTGCAAGTGATCTTGAGGCTGAACAGGTGAAAGTAAATTCACTAAC 1457

x    643 TCACATGGTGGTGGTTCATTGTTGATGAATCTAGTGAGATCGATCACGCAACTCGCTGCTTTGGAAGA  702
            ||| ||||| |||| || ||| ||||| |||||| |||| ||||| ||| ||||||||
b   1458 TCACATGGTGGTCATTGTTGATGAAAACAGTGGTGAGAGTGCTACAGCTATCCTAGAAGA 1517

y    703 ACAACTTAAGGTATTGGGAGATGCGATGGGCAAACATCGTAGATGGACAGAAGACCGCTG  762
            || | |||| ||| ||  ||| ||||| |||| |||| ||||| ||||| |||||||
b   1518 CCAGTTACAGAAACTTGGTGGCCTGGACAGCAGCAGTATGCCCGTTGGACTGAAGAACGCTG 1577

y    763 GGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTT  822
            |||  || |||||||| ||  ||||||||| |||| ||| || |||||||||| |||||
b   1578 GAATAGGTTACAAGAATCAATATATTGTGGCAGGAATTATTGGGAAGAACAGTGCTTGTT 1637

y    823 TAGTGCATGGCTTTCAGAAAAAAGAAGATGCAGTGAACAAGATTCAACAACTGGCTTAA  882
            || | ||| ||| ||||||||| |||||| |||| ||||| |  |||||||| |||||||
b   1638 GAAAGCTTGGTTAACCGAAAAAAGAGAGAGGCTTTAAATAAAGTCCAGACAACAACTTCAA 1697

y    883 AGATCAAAATGAAAATGTTATCAAGTCTTCAAAAACTGCCGTTTTAAAAGCGGATCT  939
            | |  ||||| ||| |||  || || ||| |||||||||| | |||| |||| |||
b  .1698 AGACCAAAAGGAACTAAGTGCAGTTGTTCAGTTGCACGTCTCGGCTGTGTAAGTGATGGGGTTGT 1754
```

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Wd nucleic - nucleic search, using sw model

Run on: April   4, 2004, 16:58:13 ; Search time 101.891 Seconds
           (without alignments)
           11443.127 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.........acctcagcactctggaagac 2101

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
             Maximum Match 100%
             Listing first 45 summaries

Database :   Issued Patents NA:*
     1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
     2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
     3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
     4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
     5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
     6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	1131	53.8	5952	4	US-09-687-875A-1	Sequence 1, Appli
2	1103.4	52.5	13977	4	US-09-484-970B-60	Sequence 60, Appl
3	866.6	41.2	19307	3	US-08-836-022A-10	Sequence 10, Appl
4	866.6	41.2	19307	3	US-09-427-048A-10	Sequence 10, Appl
5	343	16.3	6045	4	US-09-091-501B-7	Sequence 7, Appli
6	343	16.3	10320	4	US-09-091-501B-9	Sequence 9, Appli
7	200.8	9.6	3915	4	US-09-976-594-93	Sequence 93, Appl
8	79.4	3.8	200	4	US-09-091-501B-5	Sequence 5, Appli
9	78.6	3.7	200	4	US-09-091-501B-4	Sequence 4, Appli
10	78.6	3.7	200	4	US-09-091-501B-6	Sequence 6, Appli
11	76.6	3.6	7218	4	US-06-232-463-14	Sequence 14, Appl
12	63.6	3.0	238	4	US-09-687-875A-13	Sequence 13, Appl
13	46.2	2.2	505	4	US-09-622-976-15639	Sequence 15639, A
14	44		1230025	4	US-09-198-452A-1	Sequence 1, Appli
15	43.4	2.1	1179	4	US-09-107-532A-1186	Sequence 1186, Ap
16	40.4	1.9	832	4	US-09-622-976-2813	Sequence 2813, Ap
17	40.4	1.9	2223	1	US-08-257-073-4	Sequence 4, Appli
18	39.4	1.9	193303	4	US-09-497-855A-37	Sequence 37, Appl
19	39.4	1.9	193303	4	US-09-497-855A-44	Sequence 44, Appl
20	39.2	1.9	16995	4	US-09-961-527-82	Sequence 82, Appl
21	38.6	1.8	1751	4	US-09-620-312D-847	Sequence 847, App
22	38.6	1.8	1995	4	US-08-425-069-3	Sequence 3, Appli
23	38.6	1.8	1995	2	US-08-317-844B-3	Sequence 3, Appli
24	38.2	1.8	1131	6	5180810-3	Patent No. 5180810
25	38.2	1.8	1784	6	5180810-2	Patent No. 5180810
26	38.2	1.8	1848	4	US-09-134-001C-447	Sequence 447, App
27	38.2	1.8	4929	4	US-09-620-312D-674	Sequence 674, App
c 28		28			US-09-247-155-76	Sequence 76, Appl
c 29		37.4	1.8		US-08-985-335-4	Sequence 4, Appli
c 30		37.4	1.8		US-09-410-372-4	Sequence 4, Appli
31		37.2	1.8		US-09-092-218-1	Sequence 1, Appli
32		37.2	1.8		US-09-336-115C-5	Sequence 5, Appli
c 33		37.2	1.8		US-08-961-527-212	Sequence 212, App
34		36.8	1.8		5210183-1	Patent No. 5210183
35		36.6	1.7		US-08-887-534A-22	Sequence 22, Appl
36		36.6	1.7		US-09-527-431-22	Sequence 22, Appl
37		36.6	1.7		US-08-956-171B-263	Sequence 263, App
38		36.2	1.7		US-09-007-005-17	Sequence 17, Appl
39		36.2	1.7		US-09-244-796-17	Sequence 17, Appl
40		36.2	1.7		US-09-014-969-14	Sequence 14, Appl
c 41		36.2	1.7		US-10-204-708-23	Sequence 23, Appl
42		36.2	1.7		US-09-426-290-1	Sequence 1, Appli
43		36	1.7		US-09-385-982-236	Sequence 236, App
c 44		36	1.7		US-08-956-171B-832	Sequence 832, App
45		36	1.7		US-08-630-915A-193	Sequence 193, App

ALIGNMENTS

```
RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
;  APPLICANT: Xiao, Xiao
;  APPLICANT: Liu, Paul
;  TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PEI
;  FILE REFERENCE: 00792
;  CURRENT APPLICATION NUMBER: US/09/687,875A
;  CURRENT FILING DATE:  2000-10-13
;  PRIOR APPLICATION NUMBER: 60/158,868
;  PRIOR FILING DATE: 1999-10-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;  LENGTH: 5952
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (2897)..(2898)
;  OTHER INFORMATION: S4 junction site
;  NAME/KEY: misc_feature
;  LOCATION: (3198)..(3199)
;  OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match             53.8%;  Score 1131;  DB 4;  Length 5952;
Best Local Similarity   68.6%;  Pred. No. 0;
Matches 2101;  Conservative  0;  Mismatches  0;  Indels 960;  Gaps  1;

Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCTACACGGAGCCCATT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    891 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCTACACGGAGCCCATT 950

Qy     61 TCCTTCACAGCATTGGAGCTCCTGAGACAAGCAGTCATTGGCAGTTCATTGATGGAGAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    951 TCCTTCACAGCATTGGAGCTCCTGAGACAAGCAGTCATTGGCAGTTCATTGATGGAGAG 1010

Qy    121 TGAAGTAAACTGGACCGTTATCAAACAGCTTTAGAGAAGAAGATTATTCGTGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1011 TGAAGTAAACTGGACCGTTATCAAACAGCTTTAGAGAAGAAGATTATTCGTGCTTCTTTC 1070

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1130

Qy    241 CCAGTTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCACAGGGCCGGGTTGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1131 CCAGTTCATACTCATGAGGGTACATGATGATTGACAGCCCATCAGGGCCGGGGTGG 1190

Qy   301 TAATATTCTACAATTGGGAGGTAAGCTGATTGGAACAGGAAAAATTATCAGAGATGAAGA 360
         |||||||| |||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  1191 TAATATTCACAATTGGGAAGGTAAGCTGATTGGAACAGGAAAAATTATCAGAGATGAAGA 1250

Qy   361 AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1251 AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1310

Qy   421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1311 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1370

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAATGAGGAGGAGA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1371 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAATGAGGAGGAGA 1430

Qy   541 GCCTCTTGGACCTGATCTTGAGAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1431 GCCTCTTGGACCTGATCTTGAGAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 1490

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTGTGGTGGTAGT 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1491 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTGTGGTGGTAGT 1550

Qy   661 TGATGAATCTAGTGGACAGTCACGCCAACTGCTCTGCTTTTGGAAGAACAACTTAAGGTATTGGG 720
         |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1551 TGATGAATCTAGTGGAGAGTCACGCCAACTGCTCTGCTTTTGGAAGAACAACTTAAGGTATTGGG 1610

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1611 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1670

Qy   781 CCTTCTCAAATGGCCAACGTCTTACTGAAGACCAGTGCCTTTTTAGTGCATGGCTTTCAGA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1671 CCTTCTCAAATGGCCAACGTCTTACTGAAGACCAGTGCCTTTTTAGTGCATGGCTTTCAGA 1730

Qy   841 AAAAGAGAGACGAGTCAGTGAACAAGATCCACACAACTGGCTTTAAAGATCAAATGAAAATGTT 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1731 AAAAGAGAGACGAGTCAGTGAACAAGATCCACACAACTGGCTTTAAAGATCAAATGAAAATGTT 1790

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGATCTAGAAAAGAAATAGCAAATCCAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1791 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGATCTAGAAAAGAAATAGCAAATCCAT 1850

Qy   961 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACACTGAGCAATAAGTAAGTCAGTGAC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1851 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACACTGAGCAATAAGTAAGTCAGTGAC 1910

Qy  1021 CCAGAAGACGGAAGCAGTTAGAAGATCTGAGCTCTGAGTGTTGGGATAATTTAGTCCAAA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1911 CCAGAAGACGGAAGCAGTTAGAAGATCTGAGCTCTGAGTGTTGGGATAATTTAGTCCAAA 1970

Qy  1081 ACTTGAAAAGAGTACAGCACAG--------- 1102
         ||||||||||||||||||||||
Db  1971 ACTTGAAAAGAGTACAGCACAGCTGGGCAAACTGAAATAGCAGTTCAAGCTAAACAACCGATGT 2030

Qy  1103 ---------- 1102
Db  2031 GGAAGAGATTTGTCTAAAGGGCAGCATTGTACAAGGAAAAAACCAGCCACTCAGCCAGT 2090

Qy  1103 ---------- 1102
Db  2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGTCGGGAGGCGGTAAACCGTTTACTTCAAGA 2150

Qy  1103 ---------- 1102
Db  2151 GCTGAGGGGCAAAGCACCAGCCTGACCTGGCTCCTGGACTGACCACTATTGGAGCCTCCCTAC 2210

Qy  1103 ---------- 1102
Db  2211 TCAGACTGTTACTCGTGGTGACAACCTGTGTTGGTTACTAAGGAAACTGCCATCTCAAACT 2270
```

```
Qy  1103 ---------- 1102
Db  2271 AGAAATGCCATCTTCCTTGTGATGTTGGAGGTACCTGCTCTGCTCGCAGATTCAACCGGGCTTG 2330

Qy  1103 ---------- 1102
Db  2331 GACAGAACTTACCGACTGGCTTCTCTGCTGATCAAGTTATAAAATCACAGAGGGTGAT 2390

Qy  1103 ---------- 1102
Db  2391 GGTGGGTGACCTTGAGGATATCAACGAGATCATCAAGCAGAGGCAACAATGCAGGA 2450

Qy  1103 ---------- 1102
Db  2451 TTTGGACAGAGGCGTCCCCAGTTGGAAGAACTCATTACCGCTGCCAAAATTGAAAAA 2510

Qy  1103 ---------- 1102
Db  2511 CAAGACCAGCAATCAAGAGGCTAGAACAATCATTACGGATGAATGCGATGAAAGAATTCAGAA 2570

Qy  1103 ---------- 1102
Db  2571 TCAGTGGATGAAGTACAAGAGCAGTTGTGAATTGAGCAGTGAACAACAGTTGAATGAAATGTT 2630

Qy  1103 ---------- 1102
Db  2631 AAAGGATTCAACAATGGCTGGAAGCTAAGGAGAGAGCTGAGCAGTCTTAGGACAGGC 2690

Qy  1103 ---------- 1102
Db  2691 CAGAGCCAAGTTGAGTCATGGAAGGAGGAGGGTCCCTATACAGTAGATGATGCAATCAATCCAAAGAA 2750

Qy  1103 ---------- 1102
Db  2751 AATCACAGAAACCAAGGACAGTTGGCCCAAGAGACCTCCGGCCAGTGGCAGACAAATGTAGATGT 2810

Qy  1103 ---------- 1102
Db  2811 GGCAAATGACTTGGCCCTGAAACTTCTCCCGGGATTATTCTGCAGATGACTCCAGAAAAGT 2870

Qy  1103 ---------- 1102
Db  2871 CCACATGATAACAGAGAATATCAATGCCTCTTGGAGAAGCATTCATAAAAGGGGTGAGTGA 2930

Qy  1103 ---------ACTTATAGATTACTGCAACAGTTCCCCTGACCTGGA 1140
                  |||||||||||||||||||||||||||||||||||
Db  2931 GCGAGAGGCTGCTTGGAGAAAACTCATAGATTACTGCAACAGTTCCCCTGACCTGGA 2990

Qy  1141 AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2991 AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 3050

Qy  1201 CCGTAAGGAAAGGCTCCTAGAAGAGACTCCAGGAGCTAAAAGAGCTGATGAAAACAATGGCA 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3051 CCGTAAGGAAAGGCTCCTAGAAGAGACTCCAGGAGCTAAAAGAGCTGATGAAAACAATGGCA 3110

Qy  1261 AGACCTCCAAGTGAAGTTGAAATTGAAGCTCACACAGATGTTATCACAACCTGGATGAAAAACAG 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3111 AGACCTCCAAGTGAAGTTGAAATTGAAGCTCACACAGATGTTATCACAACCTGGATGAAAAACAG 3170

Qy  1321 CCAAAAAATCCTGAGATCCCTGGAAGTCCGATGAGTCAGTCGTTCACAAAGACGTTT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3171 CCAAAAAATCCTGAGATCCCTGGAAGTCCGATGAGTCAGTCGTTCACAAAGACGTTT 3230

Qy  1381 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTTCAACATTAGGTCCCA 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3231 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTTCAACATTAGGTCCCA 3290

Qy  1441 TTTGGAGGCCAGTTCTGACCAGTTCTGACCAGTGGAAGGCCGTCTGACCCTTTCTCTGCAGGAACTTCTGGT 1500
         ||||||||||||||||||||||                     |||||||||||||||||||||||||||||
Db  3291 TTTGGAGGCCAGTTCTGACCAGTGGAAGGCCGTCTGACCCTTTCTCTGCAGGAACTTCTGGT 3350
```

```
Matches 1104;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy     1 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy    61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1218

Qy   121 TGAAGTAAACCTGACCGGTATCAAACAGCTTAGAGAGAAGAAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 TGAAGTAAACCTGACCGGTATCAAACAGCTTAGAGAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACCAAGGAGACAGATTTCTAATGATGTGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 TGCTGAGGACACATTGCAAGCACCAAGGAGACAGATTTCTAATGATGTGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATATGACTACTCATGAGGGGTACATGATGATTGACAGCCATCAGCCCGGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTCATATGACTACTCATGAGGGGTACATGATGATTGACAGCCATCAGCCCGGGTTGG 1398

Qy   301 TAATATTCACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1399 TAATATTCACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCACATGGAAAACAAATTCACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1519 TAGCACATGGAAAACAAATTCACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAGAAGACAAGAGAAAATGGAGGAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579 GAAAGAGTTGAATGACTGGCTAACAAGAAGACAAGAGAAAATGGAGGAAGA 1638

Qy   541 GCCTCCTTGGACCTGATCTTGAAGACCTAAAACGCCAGTACACAACAACATAAGGTGCTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639 GCCTCCTTGGACCTGATCTTGAAGACCTAAAACGCCAGTACACAACAACATAAGGTGCTTCA 1698

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT 1758

Qy   661 TGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTTTGGAAGAACAACTAAGGTATTGGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1759 TGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTTTGGAAGAACAACTAAGGTATTGGG 1818

Qy   721 AGATCGATGGGCAAACATCTGTAGATGACGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819 AGATCGATGGGCAAACATCTGTAGATGACGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   841 AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAAATGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1939 AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAAATGTT 1998

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTCAAAAACGCAATCCAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1999 ATCAAGTCTTCAAAAACTGGCCGTTTCAAAAACGCAATCCAT 2058

Qy   961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2059 GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy  1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATTAATTTAGTCCAAAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATTAATTTAGTCCAAAA 2178


Qy  1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCGGCAGGCACCTATTGGAGGCGACTTTCC 3410

Qy  1561 AGCAGTTCAGAAGCAGCAGACGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3411 AGCAGTTCAGAAGCAGCAGACGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 3470

Qy  1621 ACCTGTAATCATGAGTACTCTTGAGACTGTAGACTGTACAGAATATTTCGACAGAGCAGCCTTTGGA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3471 ACCTGTAATCATGAGTACTCTTGAGACTGTAGACTGTACAGAATATTTCGACAGAGCAGCCTTTGGA 3530

Qy  1681 AGGACTAGAGAAAACTCTACCAGAGAGCTGCCTCCTGAGAGAGGCCAGAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3531 AGGACTAGAGAAAACTCTACCAGAGAGCTGCCTCCTGAGAGAGGCCAGAA 3590

Qy  1741 TGTCACTCGGCTTCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3591 TGTCACTCGGCTTCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3650

Qy  1801 CCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3651 CCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 3710

Qy  1861 TCAAGAGGGCACCGATGAGCTGGACCTGCAAGCTGACCCTCAAGAGTCAAGGGATC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3711 TCAAGAGGGCACCGATGAGCTGGACCTGCAAGCTGACCCTCAAGAGTCAAGGGATC 3770

Qy  1921 CTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTGAGAAAGTCAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3771 CTGGCAGCCCGTGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTGAGAAAGTCAA 3830

Qy  1981 GGCACTTCGAGGGAGAAATTGCGCCTCTGAAAGAGAAGAAGTGAGCCACGTCAATGACCTTGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3831 GGCACTTCGAGGGAGAAATTGCGCCTCTGAAAGAGAAGAAGTGAGCCACGTCAATGACCTTGC 3890

Qy  2041 TCGGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCTGGAAGA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3891 TCGGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCTGGAAGA 3950

Qy  2101 C 2101

Db  3951 C 3951
```

3SULT 2
5-09-484-970B-60
Sequence 60, Application US/09484970B
Patent No. 6426186
GENERAL INFORMATION:
APPLICANT: Jones, Karen A.
APPLICANT: Volkmuth, Wayne
APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PERL Program
SEQ ID NO 60
LENGTH: 13977
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
NAME/KEY: unsure
LOCATION: 11721-11761, 12294, 13969
OTHER INFORMATION: a, t, c, g, or other
5-09-484-970B-60

Query Match   52.5%; Score 1103.4; DB 4; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;

Qy   241 CCAGTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGCCGGGTTGG 300
Db 13328 ACAGTTTCATGTCATGAGGGATTCATGATGGATCTGACATCTCATCAAGACTTGTTGG 13269

Qy   301 TAATATTCTACAATTGGGAGTAAGCTAAGCTGATTGGACAGGAAAAATTATCAGAAGATGAAGA 360
Db 13268 TAATGTTCTACAGTTAGGAGTCAACTAGTTGGAAAAGGAAATTATCAGAAGATGAAGA 13209

Qy   361 AACTGAAGTACAAGAGCAGATGAATCCTCAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 13208 AGCTGAAGTGCAAGAACAAATGAATCCTCAAATTCAAGATGGGAATGTCTCAGGGTAGC 13149

Qy   421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 480
Db 13148 TAGCATGGAAAAACAAAGCAAATTACACAAGTTCTAATGATCTCCAGAATCAGAAATT 13089

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGGACAACAGGAAAATGGGAGGAGA 540
Db 13088 AAAAGAACTAGATGACTGGTTAACAAAAACTGAAGAGAAACTAAGAGAGAAATGGGAGGAAGA 13029

Qy   541 GCCTCTTGGACCTGACTCTTGAAGACCTAAAACGCCAAGTACAACACACATAAGGTGCTTCA 600
Db 13028 GCCCTTTGGACCCTGACTCTTGAAGATCTAAAATGCCAAGTACAACACACATAAGGTGCTTCA 12969

Qy   601 AGAAGATCTAGAACAAGAAGAAGTCAGGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT 660
Db 12968 AGAAGATCTAGAACAAGACAGAGAGGTCAGGGTCAACTCGCTCACTCACATGGTAGTAGTGGT 12909

Qy   661 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTGGAAGAACAACTTAAGGTATTGGG 720
Db 12908 TGATGAATCCAGCGGGTGATCATGCAACTGCTGTTGGAAGAACAACTTAAGGTACTGGG 12849

Qy   721 AGATCGATGGGCAAAACATCGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 12848 AGATCGATGGGCAAAATATCTGCAGATGGACTGGACCGCTGGGATTGTTTTTACAAGATAT 12789

Qy   781 CCCTTCAAATGCACGTCTTACTGAAGACACAGTGCCCTTTTAGTGCATGGCTTTCAGA 840
Db 12788 TCTTCTAAAATGGCAGCCAGTTTACTGAAGAACAGCAGTGCCCTTTTAGTACATGGGCTTTCAGA 12729

Qy   841 AAAAGAAGATGCAGTGAACAAGATCACAAACTGGCTTAAAGATCAAAATCAAAATGTT 900
Db 12728 AAAAGAAGATGCAATGAAGATGAACATTCAGACACAGTGCTTTAAAGATCAAAATGAAATGAT 12669

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 12668 GTCAAGTCTTCACAAAATATCTACTTTAAAATAGCATCAGAAAGAAAAAGAAAGCAACCAT 12609

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCGGACTGAAGAGTAAGTCAGTGAC 1020
Db 12608 GGAAAAACTAAGTTCACTCAATCAAGATCTACTTTCGGCACTGAAAAATAAGTCAGTGAC 12549

Qy 1021 CCAGAAGACGGAAGCATGGCATGGCTGGATAACTTTGCCCGGTGTGTTGGGATAATTTAGTCCAAAA 1080
Db 12548 TCAAAAGATGAAAATCTGGATGGAAAACTTTGCACAACGTTGGGACAATTTAACCCAAAA 12489

Qy 1081 ACTTGAAAAGAGTAGCAGCACAGACT 1105
Db 12488 ACTTGAAAAGAGTTCAGCACAAATT 12464

RESULT 4
US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
GENERAL INFORMATION:
  APPLICANT: Trustees of the University of Pennsylvania
  ;   Wilson, James M.
  ;   Fisher, Krishna J.
  ;   Chen, Shu-Jen
  ;   Weitzman, Matthew
  TITLE OF INVENTION: Improved Adenovirus Virus and
                      Methods of Use Thereof

RESULT 3
JS-08-836-022A-10/c
; Sequence 10, Application US/0836022A
; Patent No. 6001557
GENERAL INFORMATION:
  APPLICANT: Trustees of the University of Pennsylvania
  APPLICANT: Wilson, James M.
  APPLICANT: Fisher, Krishna J.
  APPLICANT: Chen, Shu-Jen
  APPLICANT: Weitzman, Matthew
  TITLE OF INVENTION: Improved Adenovirus Virus and
  NUMBER OF SEQUENCES: 10
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Howson and Howson
    STREET: Spring House Corporate Cntr, P O Box 457
    CITY: Spring House
    STATE: Pennsylvania
    COUNTRY: USA
    ZIP: 19477
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/836,022A
    FILING DATE:
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/331,381
    FILING DATE: 28-OCT-1994
  ATTORNEY/AGENT INFORMATION:
    NAME: Bak, Mary E.
    REGISTRATION NUMBER: 31,215
    REFERENCE/DOCKET NUMBER: GNVPN.008PCT
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 215-540-9200
    TELEFAX: 215-540-5818
  INFORMATION FOR SEQ ID NO: 10:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 19307 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: unknown
  MOLECULE TYPE: cDNA
JS-08-836-022A-10

Query Match   41.2%;   Score 866.6;   DB 3;   Length 19307;
Best Local Similarity   86.5%;   Pred. No. 4.1e-254;
Matches 956;   Conservative   0;   Mismatches 149;   Indels   0;   Gaps   0;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
Db 13568 GAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACCTCTGATTCCACACAGCGCCCCTA 13509

Qy   61 TCCTTCACAGCAGCTGGAGAGTCCTGAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG 120
Db 13508 TCCTTCACAGCAGCTGGAGAGTCCCGAGACAAGCAAGTCACAGTTCATTGATGGAGAC 13449

Qy 121 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db 13448 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAGTACTTTCATGCCTTCTTTC 13389

Qy 181 TGTGAGGACACATTGCAAGCACAAGGACAGAGATTTCTAATGATGTGGAGTGGTGAAAGA 240
Db 13388 TGCCGAGGATACATTGCCGAGCACAAGGAGCACAAGAGAGAGATTTCAAATGATGTTGAAGAAGTCAAAGA 13329

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
    ADDRESSEE:.Howson and Howson
    STREET: Spring House Corporate Cntr, P O Box 457
    CITY: Spring House
    STATE: Pennsylvania
    COUNTRY: USA
    ZIP: 19477
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/427,048A
    FILING DATE: 21-Oct-1999
    CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/836,022
    FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
    NAME: Bak, Mary E.
    REGISTRATION NUMBER: 31,215
    REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
    TELEPHONE: 215-540-9200
    TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
    LENGTH: 19307 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
S-09-427-048A-10

Query Match   41.2%;   Score 866.6;   DB 3;   Length 19307;
Best Local Similarity 86.5%;   Pred. No. 4.1e-254;   Indels 0;   Gaps 0;
Matches 956;   Conservative 0;   Mismatches 149;

```
Qy      1 GAGCTATGCCTACACAGGCTGCTATGTCACCACCTCGACCCTACACGAGCCCATT 60
            ||||||||||| |||||||||||||||||| || ||||||| ||||||||||||||
Db  13568 GAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACCTCGATTCCACACAGCCCCTA 13509

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 120
            |||||||||||||||||||||| || ||||||||||||| |||||||||||||||||||
Db  13508 TCCTTCACAGCATTTGGAAGCTCCCAGAGACAAGTCACTTCATTGATGGAGAC 13449

Qy    121 TGAAGTAAACCTGGACCCGTTATCAAACAGTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||| |||| ||||||||||||||||| ||||||||| |||||| ||||||||
Db  13448 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAGAGACTACTTCCATGGCTTCTTTC 13389

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
            |||||||||| ||||| |||||||||||||||||||||| ||||||||||| ||||||
Db  13388 TGCCGAGGATACATTGCGAGCACAAGGAGCACAAGGAGAGAGATTTCAAATGATGTTGAAGAAGTGAAAGA 13329

Qy    241 CCAGTTCATACTCATGAGGGGTACATGATGGGGTACAGCAGCCCATCAGGGGCCGGGGTTGG 300
            |||||||||||| ||||| || ||||||||||| |||| |||||||||||||||||||||
Db  13328 ACAGTTCATGTCATGAGGGATTCATGATGTGCAGAGTCTCATCAAGGACTTGTTGG 13269

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACACAGGAAATTATCAGAAGATGAAGA 360
            ||||||||||||||| ||| ||| ||| ||| ||||||||||||| |||||||||||||||
Db  13268 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAAATTATCAGAAGATGAAGA 13209

Qy    361 AACTGAAGTACAGAGCAGAAAGCAATTTACATGAGTTTAATGAGCCTCAGGGTAGC 420
            |||||||||||||||||||| ||||||||||||||||| ||||| ||||||||||||
Db  13208 AGCTGAAGTGCAAGAGCAAGAGAACAAATGAATCTCCTAAATTCCCAGAATCAGAAACT 13149

Qy    421 TAGCATGCAGAAAAACAAAGCAATTTACATGAGCTTTAATGAGTTTTACCACAAAGT 480
            ||||||| |||||||||||||||||||||||||||| |||||||||||||||||||
Db  13148 TAGCATGCAGAAAAACAAAGCAATTACACAAAGTTCTAATGATCTCCCAGAATCAGAAATT 13089
```

```
Qy    481 GAAAGAGTTGAATGACTGCTAACAAAAACAGAGAAGAACAAGGAAAATGAGGAGAA 540
            ||||||| ||||||||| ||||||||||||| ||||||||| ||||||||||||||
Db  13088 AAAAGAACTAGATGACTGGTTAACAAAAACTGAGAGAGAACTAAGAAAATGGAGGAAGA 13029

Qy    541 GCCTCCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
            ||||| |||||||||||||||||||| |||||| ||||||| |||| |||||||||||||
Db  13028 GCCCTTGGACCTGATCTTGAAGATCTAAAATGCCAAGTCAAGTACAACATAAGGTGCTTCA 12969

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGTGGTAGT 660
            |||||||||||||| || ||||| ||||||||||| ||| || ||| ||||| ||||||
Db  12968 AGAGATCTAGAAACAAGAGGCAGGTCAGGGCCACTCGGCTCACTCACATGGTGGTAGTGGT 12909

Qy    661 TGATGAATCTAGTGAGAACAAGTCCTGCTTGGAAGAACAACTAAGGTATTGGG 720
            |||||| ||||| ||||||| || |||||| |||||||||| || |||||||||
Db  12908 TGATGAATCCAGTGGATCATGCAACAGCTGCTTTGGAAGAACAACTTAAGGTACTGGG 12849

Qy    721 AGATCGATGGGCAAACATCTGTAGATGGCAGAGAGACCCGTGGGGTTCTTTTTACAAGACAT 780
            ||||||||| |||||||||||| || || |||||| || ||||||||||| |||||||||
Db  12848 AGATCGATGGGCAAATATCTGCAGATGGCAGAGAGACCCGTGGGGTTCTTTTTACAAGATAT 12789

Qy    781 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTTGTTAGTGCATGGCTTTCAGA 840
            || ||||||| |||||||||| |||||||||||||| ||||| ||||| ||||||||||||
Db  12788 TCTTCTAAAAATGGCAGCATTTACTGAAGAACACAGTGCCTTTTTTAGTACATGGCTTTCAGA 12729

Qy    841 AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 900
            |||||||||| |||| ||||||||| ||| || ||||| ||||| |||||||| ||| |
Db  12728 AAAAGAAGATGCAATGAAGAACATTCAGACAAGCAGTGGCTTTAAAGATCGGATCAAAATGAAATGAT 12669

Qy    901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCCGTTTAGAAGAAGAAAAAGCAATCCAT 960
            ||||| |||||| |||||| || ||| |||| |||| |||| || ||||| ||||||||
Db  12668 GTCAAGTCTTCACAAAATACTACTTAAAAAATAGATGAAAAATAGATCAGAAAAGATCAGAAAAGCCAACCAT 12609

Qy    961 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACAACAGATCTTAGTCAGTGAC 1020
            |||||||||||||||||||||||||| |||||| ||||||| || ||| ||||||||
Db  12608 GGAAAACTAAGTTCACTCAATCAAGATCTACTTCGGCACTGAAAAATAAGTCAGTGAC 12549

Qy   1021 CCAGAGACGGAAGCATGCTGGATAAACTTTGCCCGGTGTTGGGATAAATTTAGTCCAAAA 1080
            |||||| | |||| ||||| ||||| ||| || ||| |||||||| |||||||||||||||
Db  12548 TCAAAAGATGGAAATCTGGATGAAATCTTGCACAACGTTGGCACATTTAACCCAAAA 12489

Qy   1081 ACTTGAAAAGAGTACAGCACAGACT 1105
            ||||||||||||||| ||||||||||
Db  12488 ACTTGAAAAGAGTTCAGCACAAATT 12464
```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match   16.3%;   Score 343; DB 4; Length 6045;
Best Local Similarity   59.0%; Pred. No. 4.5e-94;
Matches 589; Conservative   0; Mismatches 410; Indels   0; Gaps   0;

Qy   104 AGTTCATTGATGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAGAAGTA 163
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   914 AGCACCGTCACTGAAGTGGACATGGATTTGGACAGCTGACATAGCGCTTAGAGAGAAGTG 973

Qy   164 TTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGAT 223
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   974 CTGACGTGGCTGCTGCTGTCCGCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 1033

Qy   224 GTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATTTGACAGCC 283
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1034 GTCGAAGAAGTCAAAGAGCAGTTTGCTACCCATGAAACTTTTATGATGGAGCTGACAGCA 1093

Qy   284 CATCAGGGCGCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAA 343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1094 CACCAGGAGCAGCGTGGGGAGCGTCCTGCAGGCTGGCAACCAGCTGACCAACAAGGGACT 1153

Qy   344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGG 403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1154 CTGTCAGAGGAAGAAGAAACAGCAGTGCAGGAGAGTTTGAGATCCTGAATGCAAGGTGG 1213

Qy   404 GAATGCCTCAGGTGGTAGCTAGCATGAACAGATCAGGAGAAAAACAAAGCAATTACATAGAGTTTAATGGAT 463
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1214 GAGGCGCTTCGGGTGGAGAGCATGGAGAGAGAGGCAGTCCCGGCTGCACGACGCTCTGATGGAG 1273

Qy   464 CTCCAGAATCAGAAACTGAAAGAGTGAATGACTGGCTAACAAAAACAGAGAAGAAACA 523
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1274 CTGCAGAAGAAACAGCTGCAGCTGCTCAAGCTGCTTCAAGCTGCAGAAGCTGCTTCAA 1333

Qy   524 AGGAAAATGGAGGAAGAGCCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAA 583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1334 CAGAAGAATGGAGAGCCCTCCGCTGGGTGATGACCTGCCCCCTGCAGAAGCTGCTTCAA 1393

Qy   584 CAACATAAGGTGCTTCAAGGATCTAGAACAAGAAGCAGTCAGGTCAATTCTCCTCACT 643
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1394 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAAATGCAGGTGAAGGTAAATTCCTTAACT 1453

Qy   644 CACATGGTTGGTGGTAGTTGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTGGAAGAA 703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1454 CACATGGTTGTGATTGTGATGAAAACAGTGGGACGAGTGCCACAGTCTCTTCTGGAAGAT 1513

Qy   704 CAACTTAAGGTATTGGGATCGCATGGGCAAACATCTGAGATGACAGCAGAAGACCGCTGG 763
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1514 CAGTTACAGAAGAACTGGGTGAGCCTGCTGACTGTGATCCGGCGTCTGTATTGGAGGAAGACATGGAA 1573

Qy   764 GTTCTTTTACAAGACATCCTTCTCAAATGGCAAGTCTTACTGAAGAACAGTGCCTTTTT 823
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1574 AACAGGTTGCAAGAAATCAGTATTCAGTGATTCTGTGGCAGGAATTATTGGGAACAGACTGTTG 1633

Qy   824 AGTGCATGGCTTTCAGAAGAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAA 883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1634 GAGGCTTGGCTCACGGAAAAGAAGGAGAGCCTTTGAATAAAGTTCAAACCAGCAACTTTAAA 1693

Qy   884 GATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGGTTTTAAAAGCTGAGA 943
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1694 GACCAAGGAACTAAGTGTCAGTGCTTCAGTGTCGGCGTCTCGGCGTATATTGGCTCTGGCTATAGGGAAGACATGGAA 1753

Qy   944 AAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTG 1003
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1754 ATGAAGAGGCAGACTCTGATCAACTGAGTGAGATTGGGCCAATTACTC 1813

Qy 1004 AAGAATAAGTCAGTGACCCAGAGACACGGAAGCATGGCTGGTGGATAACTTTGCCCGGTGTTGTTGG 1063

Db 1814 AGTAATCCAAGGCATCTAAGAGATGAACAGTGAACAGTACAGCACAG 1873
             |||||||||||||||||||||||||||||||||||||||||||||
Qy 1064 GATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACAG 1102
             ||||||||||||||||||||||||||||||||||||
Db 1874 GATTCTCTGGTTCAGAGACTCGAAGACTCTTCTACCAG 1912

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match   16.3%;   Score 343; DB 4; Length 10320;
Best Local Similarity   59.0%; Pred. No. 6.6e-94;
Matches 589; Conservative   0; Mismatches 410; Indels   0; Gaps   0;

Qy   104 AGTTCATTGATGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAGAAGTA 163
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   914 AGCACCGTCACTGAAGTGGACATGGATTTGGACAGCTGACATAGCGCTTAGAGAGAAGTG 973

Qy   164 TTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGAT 223
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   974 CTGACGTGGCTGCTGCTGTCCGCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGAT 1033

Qy   224 GTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATTTGACAGCC 283
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1034 GTCGAAGAAGTCAAAGAGCAGTTTGCTACCCATGAAACTTTTATGATGGAGCTGACAGCA 1093

Qy   284 CATCAGGGCGCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGACAAGGGAAA 343
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1094 CACCAGGAGCAGCGTGGGGAGCGTCCTGCAGGCTGGCAACCAGCTGACCAACAAGGGACT 1153

Qy   344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGG 403
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1154 CTGTCAGAGGAGGAGGAGAAACAGCAGTGCAGGAGAGTTTGAGATCCTGAATGCAAGGTGG 1213

Qy   404 GAATGCCTCAGGTGGTAGCTAGCATGAACAGATCAGGAGAAAAACAAAGCAATTACATAGAGTTTAATGGAT 463
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1214 GAGGCGCTTCGGGTGGAGAGCATGGAGAGAGAGGCAGTCCCGGCTGCACGACGCTCTGATGGAG 1273

```
y    464 CTCCAGAATCAGAAACTGAAAGAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAGAACA 523
          ||||||  |||||||   |||        ||||||   |                      ||
b   1274 CTGCAGAGAAACAGCTGCAGCAGCTTCAAGCTGCCTTCAGCTCCCTCACAGAGAGCGGCATT 1333

y    524 AGGAAATGGAGGAAGAGCCCTTCTTTGGACCTGATCTTGAGACGCCAAGTACAA 583
          |||        ||||      |||  || ||  |  ||| || |  |
b   1334 CAGAAGAATGCTGGTGATGCCGCGCTGGTGATGACCTGGAGAGCTGCTTCAA 1393

y    584 CAACATAAGGTGCTTCAAGAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACT 643
          |||||| || || ||||||||||||| |  ||||| || ||||||     |||   |||  |
b   1394 GAACATAAAGTTTGCAAAATGACCTTGAAGCTGGAGGTAAATTCCTTAACT 1453

y    644 CACATGGTGGTAGTTGATGAATCAGTGGAGATCACGCAACTGCTCGTGCTTTGTGGAAGAA 703
          |||||| ||  |||  || ||  |||||| ||  |  |   |||| ||
b   1454 CACATGGTGGTGATTGTGGATGAAAAACAGTGGGGAGAGTGCCACAGCTCTTCTGTGGAGAT 1513

y    704 CAACTTAAGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGG 763
          ||| | |||||   |  |||| | |  | |  ||  || |  ||||  |  |  ||
b   1514 CAGTTACAGAAACTGGGTGACGCGCGAGCAGCTGATCAGCTGGACTGAAGAAACGTTGG 1573

y    764 GTTCTTTACAGAACATCCTTCTCAAATGGCAACGTCTTACTGAGAACAGTGCCTTTTT 823
          ||||| |||| |     ||  || ||  ||  | | |  ||||||  |||   |
b   1574 AACAGGTTGCAAGAAATCAGTATTCTGTGGCAGGAATTATTGGAAGAGCCAGTGTCTGTTG 1633

y    824 AGTGCATGGCTTTCAGAAAGAGATGCAGTGACGAACAAGATTCACACAACTGGCTTTAAA 883
          ||||| | || ||| | ||  |  | |||  | | |  || |   |  || | |  |
b   1634 GAGGCTTGGCTGCACCGAAAAGGAAGGAGCTTTGAATAAAGTTCAAACCAGCAACTTTAAA 1693

y    884 GATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAA 943
          ||||| | | |||  |   | || ||  |||| | ||||
b   1694 GACCAGAGGAACTAAGTGTCAGTGTCCGGCGTGTTATTGAAGGAAGACATGGAA 1753

y    944 AAGAAAAAGCAACATCCATGGGCAAACTGTATTCACTCAAACAAGATCTCTTTTCAAACACTG 1003
          ||| | | | || |  | |     ||| | |  |||||  |  |  ||| |  | | ||||||
b   1754 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGGCCAATTACTC 1813

y   1004 AAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGG 1063
          ||    ||| | | ||     ||  |  | | | |  ||    ||| | | | ||||
b   1814 AGTAATCCCAAGGCATCTAAGAAGATGAACATGACTTCTGAGGAGCTAACAGAGAGATGG 1873

y   1064 GATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACAG 1102
          ||||| ||||   ||||| |||| | || |  |
b   1874 GATTCTCTGGTTCAGAGACTCGAAGACTCTTCTAACCAG 1912
```

```
y   1387 CATGAACTTCAAGTCGAGTCGAGTGACTCGAACTTCGGGAAAAAAGTCTCTCAACATTAGTCCATTTGGA 1446
          |||||||     |   |||  || ||  ||  |||||||| |||| ||  | ||  |   |  ||| ||
b    544 CATGAATCTGTGTTGGAATGAAATAAAAAAAAAGTCTCACAACCTCCGGCTCCGCCTCCTAGA 603

y   1447 AGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGCT 1506
          ||||| ||  |||| |  |  |||  | |||||| || || ||||| |  |||| | |
b    604 GGCCTTCTCAGACCACCAGTGAAAGCTTCAGCTCCTCCTTCAAGAGATTATTGACTGGCT 663

y   1507 ACAGTGAAAGATGATGAATTAAGCACGGCAGGCACCTATTGGAGCGACTTTCCAGCAGT 1566
          ||| |||||| ||||| || || || || |  ||||| |||| ||| |  ||  |||
b    664 CAGCCAAAAGAATGGAGTTGTCAGCTCAGCTCCCTCAGCTGCCCCTACAGGGGGATGTGGCCCTGGT 723

y   1567 TCAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGT 1626
          |||| |||| |     |||| | | | || | | |||| ||||| ||  | | ||  |
b    724 GCAACAGGAGAGGAGGACACAATGCGGCCGCCCTTTATGGAAGAGTCAAGTCTCGGGCCGCCCTA 783

y   1627 AATCATGAGTACTCTTGAGACTGTACGAATATTCTGACGAATATTCTGAAGGACT 1686
          ||| |  |||  ||| ||| | | |  |   |  |  |  || |  |  ||||||
b    784 CATCTATTCTGTCTTGGAGTCAGCTCAGGCCTTCCGTCGTCCCGATGCCACCCATTTGAGGAGTT 843

y   1687 AGAGAAACTTACCAGGAGCCCAGAGCTGCCTCCTGGAGGAGAGCCCAGAATGTCAC 1746
          || |||| | |  || |  | || |   | | |   | |||||  |||||   |
b    844 AGAGAGCCTTCATTCTGAGCAAGATACCTCCCGAAACAGCGGATCCAGAATCTCAG 903

y   1747 TCGGCTTCTACGAAAGCACGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCA 1806
          | || || || ||| ||  || || ||  | | |  || | || ||  |  |||||
b    904 CCGCTTTGTATGGAAGCAGCGGCGACGGTGGCCAGTGAACTGTGGGAGAGTTGACGCCCG 963

y   1807 CTCCGCTGACTGGCAGAGAAAATAGATGAAACCCTGAAAAGACTCCAGGAACTTCAAGA 1866
          |  | | |   || || ||  |  |   ||| ||||  ||  || |  || || |||
b    964 CTGTGTGGACCAGCCACCGCACATTGAGCCGGACTCTGGAGCGAGCTCTTGGAGATTCAGGG 1023

y   1867 GGCCACGGATGAGCTGGACCTCAAGCTGAGGTGATCAAGGGATCCTGCA 1926
          || | |||| ||    |  | |||||   |   || ||  ||| |  |||
b   1024 G--CATGGAGAACTAAGCACTACTCTGAGCTCCCGAGCTCAAGATACCTGAGAAGTCCAGCCACTTGGA 1081

y   1927 GCCCGTGGGCGAATCCTCCATTGATGACTTGCCCGGTGTTGG 1986
          |||  | |||||||||| |  | |  |||  |  | | ||||
b   1082 GCCCATTGGGAGATCTTCTTCATTGATTCACTCCCAGCACATCCCCAGACCATCAGGCTATTAAGCTGTT 1141

y   1987 TCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCA 2046
          | ||||| ||||| | ||| ||||  ||||| ||||  | || |  | || |  | | |
b   1142 CAAAGAAGAATTCTCCCCATGAAAGATGAGTAAAGTTGGTGAATGAATCTGCCCACCA 1201

y   2047 GCTTACCACTTGGGCATTCAGCTCTCACCGTATAACCTCAGCCACTCTGGAAGA 2100
          ||| |  | ||| ||| | | |  |  |  | |  | |  | || | |||||||
b   1202 ACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAATTCCCAGCCCCTGGAACA 1255
```

Query Match   9.6%; Score 200.8; DB 4; Length 3915;
Best Local Similarity   56.0%; Pred. No. 1.2e-50;
Matches 400; Conservative   0; Mismatches 312; Indels   2; Gaps   1;

; ORGANISM: Rattus sp.
JS-09-091-501B-5

Query Match   3.8%;   Score 79.4; DB 4;   Length 200;
Best Local Similarity 64.3%; Pred. No. 2.4e-14;
Matches 119; Conservative   0; Mismatches   66; Indels   0; Gaps   0;

```
Qy   565 CCTAAAACGCCAGTACACAAGTACACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGT 624
             ||  |||    ||||||||||||||||||||||| ||||| || |||| ||
Db    16 CCTGCAAAACCTGCTTGGAAGAACATAAAAAGTTTGCAAAGTGTCGAAAGCTGAGCTGAGT 75

Qy   625 CAGGGTCAATTTCTCACTCACATGGTGGTCGTAGTTGATGAATCTAGTGGAGATCACGC 684
              |   | || ||| ||||||| ||||||||||||||| |||||  ||
Db    76 GAAGGTGAATTCCTTAACTCATCATATGGTGGTGATTGTGGATGAAAACAGTGGGAGAGCGC 135

Qy   685 AACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATGCGATGGGCCAAACATCTGTAG 744
             |    ||||||||||| | |||| |||||  |||||||| | ||||||| |
Db   136 CACAGCTGTTTTGGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195

Qy   745 ATGGA 749
             |||
Db   196 CTGGA 200
```

RESULT 9
JS-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
JS-09-091-501B-4

Query Match   3.7%;   Score 78.6; DB 4;   Length 200;
Best Local Similarity 62.4%; Pred. No. 4.2e-14;
Matches 123; Conservative   0; Mismatches   74; Indels   0; Gaps   0;

```
Qy   553 TGATCTTGAAGACCTAAAACGCCAAGTACACAACATAAGGTGCTTCAAGAAGATCTAGA 612
             ||||| |  |  ||||| | ||||| |||||||||
Db     4 TGACCTGCCCTCCCTGCGAGAGCTGCTTCAGGGTAAAATGACCTTGA 63

Qy   613 ACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCTAG 672
             ||||||||||
Db    64 AGCTGAACAGGTGAAGGTGAAGGTAAATTCCTTAACTCATCGTGGTGGTGATTGTGGATGAAAACAG 123

Qy   673 TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 732
             ||||
Db   124 TGGGGAGAGTGCCACACAGCTGCCACAGCTCTTCTCGGGAGATCAGTTACAGAAACTGGGT 183

Qy   733 AAACATCTGTAGATGGA 749
             ||||
Db   184 AGCTGTATGCCGCTGGA 200
```

RESULT 10

US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match   3.7%;   Score 78.6; DB 4;   Length 200;
Best Local Similarity 62.4%; Pred. No. 4.2e-14;
Matches 123; Conservative   0; Mismatches   74; Indels   0; Gaps   0;

```
Qy   553 TGATCTTGAAGACCTAAAACGCCAAGTACACAACATAAGGTGCTTCAAGAAGATCTAGA 612
             ||||| |  |  ||||| || ||||| |||||||||
Db     4 TGATGTGAAATCTCACAAAAGCTGCTGTAGAAGAACATAAAAAGTTTGCAAAGTGATCTTGA 63

Qy   613 ACAAGAACAAGTCAGGGTCAATTCTCCTCACTCACATGGTGGTGGTAGTTGATGAATCTAG 672
             |||||||||||
Db    64 GGCTGAACAGGTGAAAGTAAAATTCACTAACTCACATGGTGTCATTGTTGTTGATGAAAACAG 123

Qy   673 TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 732
             ||||  |||||
Db   124 TGGTGAGCGGCGTACAGCTATCCTAGAAGACCAGTTACAGAAACTTGGTGAGCCGTGGAC 183

Qy   733 AAACATCTGTAGATGGA 749
             |||| |
Db   184 AGCAGTATGCCGGTTGGA 200
```

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
   APPLICATION NUMBER: US/07/935,313
   FILING DATE:
   APPLICATION NUMBER: EP 91 114 300.6
   FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
   NAME: BENT, Stephen A.
   REGISTRATION NUMBER: 29,768
   REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
   TELEPHONE: (703)836-9300
   TELEFAX: (703)683-4109
   TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
   SEQUENCE CHARACTERISTICS:
   LENGTH: 7218 base pairs
   TYPE: nucleic acid
   STRANDEDNESS: single
   TOPOLOGY: linear
IMMEDIATE SOURCE:
   CLONE: pTZgpt-F1s
S-08-232-463-14

Query Match   3.6%; Score 76.6; DB 1; Length 7218;
Best Local Similarity   6.7%; Pred. No. 2.1e-12;
Matches 28; Conservative 236; Mismatches 155; Indels   0; Gaps   0;

✓   213 TTTCTAATGATGTGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGTACATGATGG 272
         | | ||| |||||| || | || | : |:|:: : : | :
◇   1474 TATCTTATGCAAGTAGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRR 1415
✓   273 ATTTGACAGCCCATCAGGGCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTG 332
         :   : : :|: |: || :   ::|:
◇   1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355
✓   333 GAACAGGAAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAA 392
         ::   :: ::| || ::| :: || :: ||
◇   1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295
✓   393 ATTCAAGATGGGAATGCCTCAGGGTAGCATGGAAAACAAAGCAATTACATAGAG 452
         : :::: :: :| :: | : :: :::: | : :|:::
◇   1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235
✓   453 TTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAAATGACTGGCTAACAAAAACAG 512
         ::::: : :|| ::||: :: || ::::
◇   1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175
✓   513 AAGAAAGAACAAGGAAAAATGGAGGAAGAGCCTCTGGACCTGATCTTGAAGACCTAAAAC 572
         :::::::::: ||:: |::| :: | :
◇   1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115
✓   573 GCCAAGTACAACACATAAGGTGCTTCAGAAGATCTAGAACAAGAACAAGTCAGGGTC 631
         :::: :: :: ::: ::: :::: ||
◇   1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGGCAAGCTC 1056

ESULT 12
S-09-687-875A-13
; Sequence 13, Application US/09687875A
; Patent No. 654786
; GENERAL INFORMATION:
;   APPLICANT: Xiao, Xiao
;   APPLICANT: Liu, Paul
;   TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
;   FILE REFERENCE: 00792
;   CURRENT APPLICATION NUMBER: US/09/687,875A
;   CURRENT FILING DATE: 2000-10-13
;   PRIOR APPLICATION NUMBER: 60/158,868
;   PRIOR FILING DATE: 1999-10-15
;   NUMBER OF SEQ ID NOS: 22
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 238

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pXX-C2 5' junction
US-09-687-875A-13

Query Match   3.0%; Score 63.6; DB 4; Length 238;
Best Local Similarity 94.3%; Pred. No. 1.9e-09;
Matches 66; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   1348 TTCCGATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAGTGA 1407
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   169 TTCGGACGACGCAGTACTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGA 228
Qy   1408 ACTTCGGAAA 1417
          ||||||||||
Db   229 ACTTCGGAAA 238

RESULT 13
US-09-621-976-15639/c
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Jobert, S.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;   FILE REFERENCE: GENSET.054PR2
;   CURRENT APPLICATION NUMBER: US/09/621,976
;   CURRENT FILING DATE: 2000-07-21
;   NUMBER OF SEQ ID NOS: 19335
;   SOFTWARE: Patent.pm
; SEQ ID NO 15639
;   LENGTH: 505
;   TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match   2.2%; Score 46.2; DB 4; Length 505;
Best Local Similarity 17.4%; Pred. No. 0.00067;
Matches 54; Conservative 131; Mismatches 124; Indels   2; Gaps   1;

Qy   319 AAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCA 378
          ||||::|||| |||| :|||: :||||||:||:|:|: : : :||| | :| ||
Db   445 AATYAACRWAWAWGAGAAWKRWKWAGSMRACARAGTTMAWGACWAMARRGWRCARGR--T 388
Qy   379 GATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAACAAAG 438
          | || ||| || | | :|||:||: ||| :|| || | |||: || :| : : |||
Db   387 GSWGSWKGGYRMWGRGMAAKMRRMAAAGGSYCGMTSYTSGGSMTGRKGSMTKRKRWMTY 328
Qy   439 CAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTG 498
          : :::: |: |:|::: :: : :|| || |:|:|| ::: ::
Db   327 SGMWWTSYKCTKTGSKKYTGWKSKKTRWTCTSWRKYMWMMSGCWARSMKSWARSWYSMMAC 268
Qy   499 GCTAACAAAAACAGAAGAAAGAACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCT 558
          ::| : :|:::: : :::|:|:: :::: ::
Db   267 WCMMSASAYRARRSMYGARRSMMRAGAGWWRARRGKKRARGKSSMMRSKMSSMRMSAGKA 208
Qy   559 TGAAGACCTAAAACGCCAAGTACAACAGCATAAGGTGCTTCAAGAAGATCTAGAACAAGA 618
          :|:: |: ::: : :|:: :|| | :|:: ::|: :: :|:
Db   207 RMCRRWMWSCRRMSYSCMGSKCMSCRGTCAKMWRYARYAKRYASSMGKYMMGCRWCYAKC 148
Qy   619 ACAAGTCAGGG 629
          |:::|||:
Db   147 ARMYGYYRSRS 137

RESULT 14
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:

APPLICANT: Griffais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
TITLE OF INVENTION: and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (105001)..(120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (120001)..(135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (135001)..(150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (150001)..(165000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (165001)..(180000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (180001)..(195000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (195001)..(210000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (210001)..(225000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (225001)..(240000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (240001)..(255000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (255001)..(270000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (270001)..(285000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (285001)..(300000)
OTHER INFORMATION: n=a or c or g or t

NAME/KEY: misc_feature
LOCATION: (300001)..(315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (315001)..(330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

GENERAL INFORMATION:
    APPLICANT: Lynn A Doucette-Stamm and David Bush
    TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
        ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
    ADDRESSEE: GENOME THERAPEUTICS CORPORATION
    STREET: 100 Beaver Street
    CITY: Waltham
    STATE: Massachusetts
    COUNTRY: USA
    ZIP: 02354
COMPUTER READABLE FORM:
    MEDIUM TYPE: CD/ROM ISO9660
    COMPUTER: PC
    OPERATING SYSTEM: <Unknown>
    SOFTWARE: ASCII
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/107,532A
    FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 60/085,598
    FILING DATE: 14 May 1998
    APPLICATION NUMBER: 60/051571
    FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
    NAME: Ariniello, Pamela Deneke
    REGISTRATION NUMBER: 40,489
    REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
    TELEPHONE: (781)893-5007
    TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1186:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 1179 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: double
        TOPOLOGY: circular
    MOLECULE TYPE: DNA (genomic)
    HYPOTHETICAL: NO
    ANTI-SENSE: NO
    ORIGINAL SOURCE:
        ORGANISM: Enterococcus faecium
    FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (B) LOCATION 1..1179
    SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186

Query Match   2.1%;   Score 43.4; DB 4; Length 1179;
Best Local Similarity 50.7%;   Pred. No. 0.0087;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy   471 ATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAA 530
         | ||| |||| |||||||| ||| |||| |||||||||| | ||||| |||
Db   602 AGCAAGAATAAAAGATTTGATCGATCAGACAAAGAAAAATGGAGATACGATCGGAGGAA 661

Qy   531 TGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATA 590
         ||| ||||| || |||| |||| |||||||||||||| |||||||||||||||
Db   662 TTGTAGAGTGCTCGTTGGAGGCGTTCCAGCTGGATTAGGAAGTACGTACAATGGGACA 721

Qy   591 AGGTGCTTCAAGAGATCTAGAACAAGAACAAGCAAGTCAGGGTCAATTCTCTCACTCACATGG 650
         ||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||
Db   722 CGAAGCTAGATGCCAAAATCGCACAAGCTGTGGTTAGTATCAATGCCTTTAAAGGCGTAG 781

Qy   651 TGGTGGTAGTTGATGAATCTAGTGG 675
         || |||||| ||||| ||
Db   782 AATTTGGGGTCGGATTCACTTCTG 806

Search completed: April 5, 2004, 08:24:46
Job time : 116.891 secs

LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t .
NAME/KEY: misc_feature

Query Match   2.1%;   Score 44; DB 4; Length 1230025;
Best Local Similarity 53.5%;   Pred. No. 0.7;
Matches 92; Conservative 0; Mismatches 0; Indels 80; Gaps 0;

y   342 AATTATCAGAAGAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGAT 401
         ||| |||| ||||||| ||||||
b 656477 AATTAGAAGAAAGAAGAGAGAGAAGAAGAGAAATTGAGGATATCAAAGACTTCAGATACAAAAT 656418

y   402 GGGAAATGCCTCAGGGTAGCTAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTTAATGG 461
         |||| |||||||||||| ||||
b 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCACTAGGCAAGCAATTTATGTGG 656358

y   462 ATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAAACAGA 513
         ||| || | ||||
b 656357 CAATTAAGCAGAAAAAACTAAAAGCTTCTAAAGAGACGCGCTGGGAAATAGA 656306

ESULT 15
S-09-107-532A-1186
Sequence 1186, Application US/09107532A
Patent No. 6583275

```
          Copyright    GenCore version 5.1.6
                       Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

run on:    April  5, 2004, 04:53:30 ; Search time 509.455 Seconds
                                      (without alignments)
                              15454.047 Million cell updates/sec

title:           US-09-845-416-2_COPY_900_3000
perfect score:   2101
sequence:        1 gagctatgcctacacagg........acctcagcacctgaaagac 2101

scoring table:   IDENTITY.NUC
                 Gapop 10.0 , Gapext 1.0

searched:        2470430 seqs, 1873665578 residues

total number of hits satisfying chosen parameters:        4940860

minimum DB seq length: 0
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

database :       Published Applications NA:*
              1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
              2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
              3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
              4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
              5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
              6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
              7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
              8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
              9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
             10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
             11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
             12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
             13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
             14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
             15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
             16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
             17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
             18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
sult          Query
No.  Score   Match Length DB  ID                 Description
---- -----   ----- ------ --  --                 -----------
  1  2101    100.0  4182  10  US-09-845-416-2    Sequence 2, Appli
  2  2101    100.0  5149  10  US-09-845-416-27   Sequence 27, Appl
  3  1579.8   75.2  3999  10  US-09-845-416-6    Sequence 6, Appli
  4  1579.8   75.2  4966  10  US-09-845-416-28   Sequence 28, Appl
  5  1579.8   75.2  4990  10  US-09-845-416-34   Sequence 34, Appl
  6  1443     68.7  3858  10  US-09-845-416-9    Sequence 9, Appli
  7  1443     68.7  4825  10  US-09-845-416-29   Sequence 29, Appl
  8  1443     68.7  4848  10  US-09-845-416-35   Sequence 35, Appl
  9  1443     68.7  5060  10  US-09-845-416-36   Sequence 36, Appl
 10  1293     61.5  8689  15  US-10-149-736-42   Sequence 42, Appl
 11  1114     53.0  4414  15  US-10-149-736-39   Sequence 39, Appl
 12  1103.4   52.5  5417  15  US-10-149-736-39   Sequence 39, Appl
 13  1103.4   52.5 11443  15  US-10-149-736-44   Sequence 44, Appl
 14  1103.4   52.5 12057  15  US-10-149-736-47   Sequence 47, Appl
 15  1103.4   52.5 13957   9  US-09-782-378A-22  Sequence 22, Appl
 16  1103.4   52.5 13957   9  US-09-880-107-2284 Sequence 2284, Ap
 17  1103.4   52.5 13957  15  US-10-149-736-1    Sequence 1, Appli
 18  1103.4   52.5 14069  12  US-10-342-887-434  Sequence 434, App
 19  1103.4   52.5 14082  12  US-10-342-887-981  Sequence 981, App
 20  1103.4   52.5 14082  10  US-10-341-434-108  Sequence 108, App
 21  1102     52.5  3446  10  US-09-845-416-14   Sequence 14, Appl
 22  1100.2   52.4 11058  10  US-09-845-416-1    Sequence 1, Appli
 23  1097.8   52.3  1991  10  US-09-845-416-3    Sequence 3, Appli
 24            999  47.5  2169  10  US-09-845-416-4    Sequence 4, Appli
 25            999  47.5  3531  10  US-09-845-416-10   Sequence 10, Appl
 26            999  47.5  4498  10  US-09-845-416-30   Sequence 30, Appl
 27            997  47.5  5339  15  US-10-149-736-40   Sequence 40, Appl
 28            996  47.4  5462  15  US-10-149-736-41   Sequence 41, Appl
 29         866.6  41.2 13815  15  US-10-149-736-2    Sequence 2, Appli
 30           787  37.5  3510  10  US-09-845-416-12   Sequence 12, Appl
 31           787  37.5  4476  10  US-09-845-416-31   Sequence 31, Appl
 32           777  37.0  1667  10  US-09-845-416-7    Sequence 7, Appli
 33           652  31.0  1821  10  US-09-845-416-13   Sequence 13, Appl
 34           450  21.4  1340  10  US-09-845-416-11   Sequence 11, Appl
 35           387  18.4   387  10  US-09-845-416-7    Sequence 7, Appli
 36         361.4  17.2 10302   9  US-09-782-378A-23  Sequence 23, Appl
 37         361.4  17.2 10302  15  US-10-149-736-3    Sequence 3, Appli
 38         361.4  17.2 16531  14  US-10-101-510-667  Sequence 667, App
 39           348  16.6   348  15  US-10-149-736-31   Sequence 31, Appl
 40         336.6  16.0 11096  15  US-10-149-736-4    Sequence 4, Appli
 41           333  15.8   333   9  US-09-845-416-9    Sequence 9, Appli
 42           327  15.6   327  15  US-10-149-736-8    Sequence 8, Appli
 43         322.4  15.3   333  15  US-10-149-736-10   Sequence 10, Appl
 44           265  12.6  1434  15  US-10-149-736-15   Sequence 15, Appl
 45           261  12.4   324  15  US-10-149-736-33   Sequence 33, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-845-416-2
; Sequence 2, Application US/0984416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
;   APPLICANT: XIAO, XIAO
;   TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;   TITLE OF INVENTION: THEREOF
;   FILE REFERENCE: DE1142
;   CURRENT APPLICATION NUMBER: US/09/845,416
;   CURRENT FILING DATE: 2001-04-30
;   PRIOR APPLICATION NUMBER: 60/200,777
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4182
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-845-416-2

Query Match             100.0%;  Score 2101;  DB 10;  Length 4182;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 2101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCGACCGGAGCCCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCGACCGGAGCCCATT 959

Qy    61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTTCATTGGCAGTTGATGGAGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTTCATTGGCAGTTGATGGAGAG 1019

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
```

Qy 1080 ||||||||||||||||||||||||||||||||||||||||||||||| 1139
Db   TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGTGAAAGA

Qy   241 CCAGTTTCATACTCATGAGGGGTACATGATGGGCCATCAGGGCCGGGTTGG 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1140 CCAGTTTCATACTCATGAGGGGTACATGATGGGCCATCAGGGCCGGGTTGG 1199

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTAGTTGGAACAGGAGAAAATTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA 1259

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1379

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAAGAACAAGGAAAATGGAGGAAGA 540
Db 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAAGAACAAGGAAAATGGAGGAAGA 1439

Qy   541 GCCTTCTTGGACCTGATCTTGAAGACAAGAGCCCAAGTACAACACATAAGGTGCTTCA 600
Db 1440 GCCTTCTTGGACCTGATCTTGAAGACAAGAGCCCAAGTACAACACATAAGGTGCTTCA 1499

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTTCACTCACATGGTGGTAGT 660
Db 1500 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTTCACTCACATGGTGGTAGT 1559

Qy   661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGAAGAACAACTAAGGTATTGGG 720
Db 1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGAAGAACAACTAAGGTATTGGG 1619

Qy   721 AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAGACAT 780
Db 1620 AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAGACAT 1679

Qy   781 CCTTCTCAAATGGGCAACGTCTTACTGACGAACAAGTGCCTTTTAGTGCATGCTTTCAGA 840
Db 1680 CCTTCTCAAATGGGCAACGTCTTACTGACGAACAAGTGCCTTTTAGTGCATGCTTTCAGA 1739

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACCACAACTGGCTTAAAGATCAAAATGAAATGTT 900
Db 1740 AAAAGAAGATGCAGTGAACAAGATTCACCACAACTGGCTTAAAGATCAAAATGAAATGTT 1799

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGGATCTAGAAAAGAAAAAGCAATCCAT 960
Db 1800 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGGATCTAGAAAAGAAAAAGCAATCCAT 1859

Qy   961 GGGCAAACTGTATTCACTCAAACAAGAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1020
Db 1860 GGGCAAACTGTATTCACTCAAACAAGAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1919

Qy 1021 CCAGAAGACGGAAGCACATGCAGATGCGGATGCTGGATAACTTTGCCCGGTGTGTTCCCCGTGTCCACCTGGA 1080
Db 1920 CCAGAAGACGGAAGCACGGATGCTGGATAACTTTGCCCGGTGTGTTGGGATATTTGGGATATTTGGATCCAAAA 1979

Qy 1081 ACTTGAAAAGAGTACAGCACAGACAGATTACTCATAGATTACTGCAACAGTTCCCCTGGA 1140
Db 1980 ACTTGAAAAGAGTACAGCACAGACAGATTACTCATAGATTACTGCAACAGTTCCCCTGGA 2039

Qy 1141 AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCCAATGCCAAGCTAC 1200
Db 2040 AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCCAATGCCAAGCTAC 2099

Qy 1201 CCGTAAGGAAAAGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 1260
Db 2100 CCGTAAGGAAAAGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 2159

Qy 1261 AGACCTCCAAGGTGAAATTGAAGCTCACCACAGATGTTTATCACAACCTGGATGAAAACAG 1320

Qy 2160 AGACCTCCAAGGTGAAATTGAAGCTCACCACAGATGTTTATCACAACCTGGATGAAAACAG 2219
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCAGTGCCTGTTACAAAGACGTTT 1380
Db 2220 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCAGTGCCTGTTACAAAGACGTTT 2279

Qy 1381 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAAGTCTTCAACATTAGGTCCCA 1440
Db 2280 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAAGTCTTCAACATTAGGTCCCA 2339

Qy 1441 TTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGAAGCGTCTGCACCTTTCTCTGCGAGGACTTCTGGT 1500
Db 2340 TTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCGAGGACTTCTGGT 2399

Qy 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGGACTTTCC 1560
Db 2400 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGGACTTTCC 2459

Qy 1561 AGCAGTTCAGAAGCCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA 1620
Db 2460 AGCAGTTCAGAAGCCAGCAGACGATGTACATAGGGCTTCAAGAGGGAATTGAAAACTAAAGA 2519

Qy 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTAGAAATATTTCGACAGAGCCCAGAA 1680
Db 2520 ACCTGTAATCATGAGTACTCTTGAGACTGTAGAAATATTTCGACAGAGCCCAGAA 2579

Qy 1681 AGGACTAGAGAAACTCTACCAGAGGCCCAGAGCTGAGGAGTCAATATCGAGTGGGAAAAATTGAA 1740
Db 2580 AGGACTAGAGAAACTCTACCAGAGGCCCAGAGCTGAGGAGTCAATATCGAGTGGGAAAAATTGAA 2639

Qy 1741 TGTCACTCGGCTTCTACCAGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 1800
Db 2640 TGTCACTCGGCTTCTACCAGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 2699

Qy 1801 CCTGCACTCCGCTGACTGGCGAGAGAAAAAATAGATGAGACCCTGAAAGACTCCAGGAACT 1860
Db 2700 CCTGCACTCCGCTGACTGGCGAGAGAAAAAATAGATGAGACCCTGAAAGACTCCAGGAACT 2759

Qy 1861 TCAAGAGGCCACCGATGAGCTGACCTCAAGTGCGCCAAGCTGAGGTGATCAAGGGATC 1920
Db 2760 TCAAGAGGCCACCGATGAGCTGACCTGGCGACCTGAGTGATCAAGGGATC 2819

Qy 1921 CTGGCAGCCCGTGGCGACTCTCCTCATTGGCTTCTCTTCCAAGATCACCTCGAGAAGTCAA 1980
Db 2820 CTGGCAGCCCGTGGCGACTCTCCTCATTGGCTTCTCTTCCAAGATCACCTCGAGAAAGTCAA 2879

Qy 1981 GGCACTTCGAGGAGAAATTGCGCCTTGCCTCTGAAAGAGAACGTGAGCGACCAGTCAATGACCTTGC 2040
Db 2880 GGCACTTCGAGGAGAAATTGCGCCTTGCCTCTGAAAGAGAACGTGAGCGACCAGTCAATGACCTTGC 2939

Qy 2041 TCGGCCAGCTTACCACTTTGGGCATTCAGCTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
Db 2940 TCGGCCAGCTTACCACTTTGGGCATTCAGCTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2999

Qy 2101 C 2101
Db 3000 C 3000

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 27
LENGTH: 5149
TYPE: DNA
ORGANISM: Homo sapiens
S-09-845-416-27

Query Match   100.0%; Score 2101; DB 10; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 GAGCTATGCTACACACAGCCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1657 GAGCTATGCTACACACAGCCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1716

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAG 1776

Qy  121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGCTGCTTCTTC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGCTGCTTCTTC 1836

Qy  181 TGCTGAGGACACATTGCAAGCACAAGGACACAAGGAGAGATTCTAATGATGTGGAAGTGTGAAAGA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1837 TGCTGAGGACACATTGCAAGCACAAGGACACAAGGAGAGATTCTAATGATGTGGAAGTGTGAAAGA 1896

Qy  241 CCAGTTCATACTCATGAGGGTACAGCCCATCAGGGCCGGGTTGG 300
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1897 CCAGTTCATACTCATGAGGGTACAGCCCATCAGGGCCGGGTTGG 1956

Qy  301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAAGAGATGAAGA 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAAGAGATGAAGA 2016

Qy  361 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2017 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

Qy  421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATGCATCCCAGAATCAGAAACT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2077 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATGCATCCCAGAATCAGAAACT 2136

Qy  481 GAAAGAGTTGAATGGGCAAACATGTCTAGAATGGACAAGAACAAGAGAAAAATGGAGGAGA 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2137 GAAAGAGTTGAATGGGCAAACATGTCTAGAATGGACAAGAACAAGAGAAAAATGGAGGAGA 2196

Qy  541 GCCTCTTGGACCTGATCTTGAAGACTAAAACGCCAAGTACACAACATAAGGTGCTTCA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2197 GCCTCTTGGACCTGATCTTGAAGACTAAAACGCCAAGTACACAACATAAGGTGCTTCA 2256

Qy  601 AGAAGATCTAGAACAAGAACAAGAAGTCAGGGTCAATTCTCTCACTCACTCACTTAAGGTGTGGTAGT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2257 AGAAGATCTAGAACAAGAACAAGAAGTCAGGGTCAATTCTCTCACTCACTCACTTAAGGTGTGGTAGT 2316

Qy  661 TGATGAATCTAGTGGAGATCACGCAACTGCTTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376

Qy  721 AGATCGATGGGCAAACATGGGCAACATCTGTAGATGGACAGAGAAAAAATGAGACCCTGGGTTCTTTTACAAGACAT 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2377 AGATCGATGGGCAAACATGGGCAACATCTGTAGATGGACAGAGAAAAAATGAGACCCTGGGTTCTTTTACAAGACAT 2436

Qy  781 CCTTCTCAAATGGCAAGTGCCTTTTAGTGCATGGCATGGCTTTTCAGA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 2437 CCTTCTCAAATGGCAAGTGCCTTTTAGTGCATGGCATGGCTTTTCAGA 2496

Qy  841 AAAGAAGATGCAGTGAACAAGATTCACACAACAGTCATGGCTTTAAAGATCAAAATGAATGTT 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2497 AAAGAAGATGCAGTGAACAAGATTCACACAACAGTCATGGCTTTAAAGATCAAAATGAATGTT 2556

Qy  901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAGAGAAAAAGCAATCCAT 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2557 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAGAGAAAAAGCAATCCAT 2616

Db 2557 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAGAGAAAAAGCAATCCAT 2616
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961 GGGCAAACTGTATTCACTCAAACAGAGATCTTCTTCAACACTGAAGAGATAAGTCAGTGAC 1020

Db 2617 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAGATAAGTCAGTGAC 2676
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021 CCAGAAGACGGAAGCATGGCATGCTGGATAACTTTGCCCGGTGTTGGGGATAATTAGTCCAAAA 1080

Db 2677 CCAGAAGACGGAAGCATGGCATGCTGGATAACTTTGCCCGGTGTTGGGGATAATTAGTCCAAAA 2736
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081 ACTTGAAAAGAGTACAGACACAGACTCATAGATTACTGACAGTACCCCCTGGACCTGGA 1140

Db 2737 ACTTGAAAAGAGTACAGACACAGACTCATAGATTACTGACAGTACCCCCTGGACCTGGA 2796
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1141 AAAGTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1200

Db 2797 AAAGTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 2856
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAGTGATGAAACAATGGCA 1260

Db 2857 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAGTGATGAAACAATGGCA 2916
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320

Db 2917 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2976
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1321 CCAAAAAATCCTGAGATCCTGAAGGTTCCGATGATGCAGTTCCTGTTACAAAGACGTTT 1380

Db 2977 CCAAAAAATCCTGAGATCCTGAAGGTTCCGATGATGCAGTTCCTGTTACAAAGACGTTT 3036
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1381 GGATAACATGAACTTCAAGTGGAGTGAACTGGAAGGCTCTCGCGAGGAACTTCTGGT 1440

Db 3037 GGATAACATGAACTTCAAGTGGAGTGAACTGGAAGGCTCTCGCGAGGAACTTCTGGT 3096
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1441 TTTGGAAGCCAGTTCTGACCAGTGACCAGTCCGGCAGGCACCTATTGGAGGCGACTTTC 1500

Db 3097 TTTGGAAGCCAGTTCTGACCAGTGACCAGTCCGGCAGGCACCTATTGGAGGCGACTTTC 3156
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1501 GTGGCTACAGCTCAGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560

Db 3157 GTGGCTACAGCTCAGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 3216
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1561 AGCAGTTCAGAAGCAGAAGACGATGTACATAGGGCAATTGAAAACTAAAGA 1620

Db 3217 AGCAGTTCAGAAGCAGAAGACGATGTACATAGGGCAATTGAAAACTAAAGA 3276
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACAGAATATTTCTGACAGAGCAGCCTTTGGA 1680

Db 3277 ACCTGTAATCATGAGTACTCTTGAGACTGTACAGAATATTTCTGACAGAGCAGCCTTTGGA 3336
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1681 AGGACTAGAGAAACTCTACCAGAAGCCCAGGAGGCTGCTCCTGAGGAGAGCCCAGAA 1740

Db 3337 AGGACTAGAGAAACTCTACCAGAAGCCCAGGAGGCTGCTCCTGAGGAGAGCCCAGAA 3396
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1741 TGTCACTCGGCTTCACGAAAGCAAGGCTGAGGAGTCAATACTGAAGTGGGAAAAATTGAA 1800

Db 3397 TGTCACTCGGCTTCACGAAAGCAAGGCTGAGGAGTCAATACTGAAGTGGGAAAAATTGAA 3456
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1801 CCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1860

Db 3457 CCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 3516
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1861 TCAAGAGGCCACGGATGAGCTGACCTTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1920

Db 3517 TCAAGAGGCCACGGATGAGCTGACCTTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 3576
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1921 CTGGCAGCCGCCTCGAGGAGAAATTGCGCCTCCCATTGACTCCTCCCAAGATCACCTCGAGAAAGTCAA 1980

Db 3577 CTGGCAGCCGCCTCGAGGAGAAATTGCGCCTCCCATTGACTCCTCCCAAGATCACCTCGAGAAAGTCAA 3636
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1981 GGCACTTCGAGGAGAGAAATTGCGCCTCCTGAAAGAAGACAGTGAGCCACGTCAATGACCTTGC 2040

Db 3637 GGCACTTCGAGGAGAGAAATTGCGCCTCCTGAAAGAAGACAGTGAGCCACGTCAATGACCTTGC 3696
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
y   2041  TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACGCTATAAACCTCAGCACTCTGGAAGA  2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3697  TCGCCCAGCTTACCACTTTGGGCATTCAGCTCTCACGCTATAAACCTCAGCACTCTGGAAGA  3756

y   2101  C  2101
          |
b   3757  C  3757
```

ESULT 3
S-09-845-416-6
Sequence 6, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 3999
TYPE: DNA
ORGANISM: Homo sapiens
S-09-845-416-6

```
Query Match         75.2%;  Score 1579.8;  DB 10;  Length 3999;
Best Local Similarity 87.9%;  Pred. No. 0;
Matches 1846;  Conservative 0;  Mismatches 72;  Indels 183;  Gaps 4;

Qy    1  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTGACCTGACCTGACCGGAGCCCATT  60
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  900  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCTGACCTGACCGGAGCCCATT  959

Qy   61  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGCTTCTTTC  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  960  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  1019

Qy  121  TGAAGTAAACCTGGACCCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1020  TGAAGTAAACCTGGACCCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC  1079

Qy  181  TGCTGAGGACACATTGCAAGCACAAGCAACGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1080  TGCTGAGGACACATTGCAAGCACAAGCAACGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  1139

Qy  241  CCAGTTCATACTCATGAGGGGTACATGAGGTAGTGACATTGACAGCCCATCAGGGCCGGGGTTGG  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1140  CCAGTTCATATCTCATGAGGGGTACATGAGGTAGTGACATTGACAGCCCATCAGGGCCGGGGTTGG  1199

Qy  301  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1200  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA  1259

Qy  361  AACTGAAGTACAAGAGAGCAGATGAATCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1260  AACTGAAGTACAAGAGAGCAGATGAATCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC  1319

Qy  421  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGAATCCCAGAAATCAGAAACT  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1320  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGAATCCCAGAAATCAGAAACT  1379

Qy  481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAACAAGGAAAATGGAGGAAGA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1380  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAACAAGGAAAATGGAGGAAGA  1439

Qy  541  GCCTCTTGGACCTGATCTTGAAGACCTAAAAGCCAAGTACAACAACATAAGTGCTTCA  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1440  GCCTCTTGGACCTGATCTTGAAGACCTAAAAGCCAAGTACAACAACATAAGTGCTTCA  1499

Qy  601  AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCCTCACTCACATGTGTGGTAGT  660
         ||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db 1500  AGAGATCTAGAACAAGAACAAGTCAGGTCAATTCCTCACTCACATGTGGTGGTAGT  1559

Qy  661  TGATGAATCCAGTGGATGGAGACACGCAACTCTGCTTGGAAGACAACTAAGGTATTGGG  720
         ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db 1560  TGATGAATTCAGTGGATGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG  1619

Qy  721  AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAAGACAT  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1620  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACCA  1679

Qy  781  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGAGTGCATGGCTTCAGA  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1680  GCCTGAC----------CTAGCTCCTGGACTGACAACTATTGGAGCCTCTCCTACTCA  1727

Qy  841  AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAAATGTT  900
         ||||||||| |||||||                  |||   |  |   || ||||
Db 1728  GACTGTTACTCTGGTGACACA----------ACCTGTGGTTACTAAGGAAACTGCCATCT-  1777

Qy  901  ATCAAGTCTTCAAAAACTCGCCGTTTAAAGCGGATCTAGAAAGAAGAAAAGCAATCCAT  960
         ------------------------------------------------------
Db 1778  --------------------------------------------------------  1777

Qy  961  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCACACTGAAGAATAAGTCAGTGAC  1020
         ||||                                           |
Db 1778  ------------------------------CCAAACTAGAAATGCCATCTCCTTGATGTGTGGAG------  1812

Qy 1021  CCAGAAGACCGAAGCATGGCTGGATAACTTTGCCCGTGTGTTGGGATAATTTAGTCCAAAA  1080
         ----------------------------------------------------
Db 1813  --------------------------------------------------------  1812

Qy 1081  ACTTGAAAGAGTACAGCACCAGACTCATAGATTACTGCAACAGTTCCCCCTGACCTGGA  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1813  --------GTACCTACCTCATAGATTACTGCAACAGTTCCCCTGACCTGGA  1856

Qy 1141  AAAGTTCTTGCCTGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC  1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1857  AAAGTTCTTGCCTGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC  1916

Qy 1201  CCGTAAGGAAAGGTCCTCAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCA  1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1917  CCGTAAGGAAAGGTCCTCAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCA  1976

Qy 1261  AGACCCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG  1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1977  AGACCTTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG  2036

Qy 1321  CCAAAAATCCTGACATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT  1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2037  CCAAAAATCCTGACATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT  2096

Qy 1381  GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA  1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2097  GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA  2156

Qy 1441  TTTGGAAGCCAGTTCTGACCAGTGAAGCCTCTGCACCCTTTCTCGCAGGAACTTCTGGT  1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2157  TTTGGAAGCCAGTTCTGACCAGTGAAGCCTCTGCACCCTTTCTCGCAGGAACTTCTGGT  2216

Qy 1501  GTGGCTACACTGAAAGATGAAAGATGATCCCAGAATCAGAAACT  1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2217  GTGGCTACAGTGAAAGATGAAAGATGATCCCAGAATCAGAAACT  2276

Qy 1561  AGCAGTTCAGAGACAGAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA  1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2277  AGCAGTTCAGAGACAGAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA  2336

Qy 1621  ACCTGTAATCATGAGTACTCTTGAGACTGTACGGAATATTTCTGACAGCAGCCTTTGGA  1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2337  ACCTGTAATCATGAGTACTCTTGAGACTGTACGGAATATTTCTGACAGCAGCCTTTGA  2396
```

```
1681 AGGACTAGAGAAACTTCACCAGGAGCTGCCTCCTGAGGAGAGCCCAGAA 1740
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2397 AGGACTAGAGAAACTCTACCAGGAGCCCAGGAGAGCTGCCTCCTGAGGAGAGAGAGCCCAGAA 2456

1741 TGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2457 TGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 2516

1801 CCTGCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACT 1860
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2517 CCTGCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACT 2576

1861 TCAAGAGGGCCACGGATGAGCTGGACCTCCAAGCTGCGCCAAGCTGAGTGATCAAGGGATC 1920
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2577 TCAAGAGGGCCACGGATGAGCTGGACCTCCAAGCTGCGCCAAGCTGAGTGATCAAGGGATC 2636

1921 CTGGCAGCCCGTGGGCGATCTCCTCCATTGACTGACTCTCCAAGATCACCTCGAGAAGTCAA 1980
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2637 CTGGCAGCCCGTGGGCGATCTCCTCCATTGACTGACTCTCCAAGATCACCTCGAGAAGTCAA 2696

1981 GGCACTTCGAGGAGAAATTGCGCCTTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2697 GGCACTTCGAGGAGAAATTGCGCCTTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2756

2041 TCGCCAGCTTACCAACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCACTCTGGAAGA 2100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2757 TCGCCAGCTTACCAACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCACTCTGGAAGA 2816

2101 C 2101
     |
2817 C 2817
```

```
1837 TGCTGAGGACACATTGCAGCACAAGGAGGAGATTCTAATGATGTGGAAGTGTGGTGAAAGA 1896
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 241 CCAGTTTCATACTCATGAGGGGTACATGATGAAGTTGACAGCCATCAGGCGCCGGGGTTGG 300

1897 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCATCAGGCGCCGGGGTTGG 1956
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 301 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA 360

1957 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA 2016
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 420

2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480

2077 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 2136
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAAATGAGGAGAAGA 540

2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAAATGAGGAGAAGA 2196
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 541 GCCTCTTGACCTGATCTTGAAGAACGTGAGCCAGTACAACAACATAAGGTGCTTCA 600

2197 GCCCTTTGGACCTTGGACCTGATCTTGAAGACGTCAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCCTCACCTCCACTCACATGGTGGTAGT 660

2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCCTCACCTCCACTCACATGGTGGTAGT 2316
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 661 TGATGAATCTAGTGGAGTCAGTGGAGATCGCAACTTAAGTATTGGG 720

2317 TGATGAATCTAGTGGAGTCAGTGGAGATCACCGCAACTTAAGTATTGGG 2376
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 721 AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTCTTCTTTTACAAGACAT 780

2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTCTTTTACAAGACCA 2436

 781 CCTTCTCAAATGGCAACGTCTTACTCGAAGAACAGTGCCTTTAGTGCATGGCTTTCAGA 840
     |||||||||||||         |||||||||||||||||||
2437 GCCTGAC----------CTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTCCTACTCA 2484

 841 AAAAGAAGATGCAGTGAACAAGATTCACCACAACTGGCTTAAAGATCAAAATGAAATGTT 900
                   ||||    ||||||||||        |||
2485 GACTCTTACTCTGGTGACACA-------------ACCTGTGTTACTAAGGAAACTGCCATCT- 2534

 901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAGAAAAGAAAGCAATCCAT 960
     ||||||||||||||||||||||||||||||
2535 ------------------------------------------------------------ 2534

 961 GGGCAAACTGTATTCACTCAAACAAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
     |||||||  ||||            |||||||||  |
2535 --------------CCAAACTAGAAATGCCATTCCATTCCTTGATGTTGGAG--------- 2569

1021 CCAGAAGACGGAAGCATGGCGGTGTTGCCCGGTATAATTTAGTCAAAA 1080
     |||||||||||||||
2570 ------------------------------------------------------------ 2569

1081 ACTTGAAAGAGTACAGACCACAGACTCATAGATTCCCCCTGACCAGTTCCCCTGA 1140
     ||||||||||||||||||||      ||||||||||||
2570 --------GTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA 2613

1141 AAAGTTTCTTGCCTGGCTTACAGAAGTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1200
     ||||||||||||||||||||      ||||||||||
2614 AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 2673

1201 CCGTAAGCAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCA 1260
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2674 CCGTAAGCAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCA 2733

1261 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAACCTGATGAAAACAG 1320
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2734 AGACCTCCAAGTGAAATTGAAGCTGAAAATTGAAGCTCACACAGATGTTTATCAACCTGATGAAAACAG 2793
```

Query Match   75.2%; Score 1579.8; DB 10; Length 4966;
Best Local Similarity   87.9%; Pred. No. 0;
Matches 1846; Conservative   0; Mismatches 72; Indels 183; Gaps 4;

```
   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACACTTCGACCTTCCACCGGAGCCCATT 60
     |||||||||||||||||||||||||||||||||||||||||||||||||
1657 GAGCTATGCCTACACCACCAGCTGCTTATGTCACCACCTCCTGACCCTCGACCGGAGCCCATT 1716

  61 TCCTTCACAGCATTGGAAGCTCCTGAAGACACAAGTCATTGGCAGTTCATTGATGAAGAG 120
     |||||||||||||||||||||||||||||||||||||||||||||||
1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACACAAGTCATTTGGCAGTTCATTGATGAGGAG 1776

 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1836

 181 TGCTGAGGACACACATTGCAGCACAAGGAGGAGATTCTAATGATGTGGAAGTGTGGTGAAGA 240
     ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
;  SEQ ID NO 34
;    LENGTH: 4990
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-845-416-34

Query Match            75.2%;   Score 1579.8;   DB 10;   Length 4990;
Best Local Similarity  87.9%;   Pred. No. 0;
Matches 1846;  Mismatches  0;  Conservative  0;  Indels  183;  Gaps  4;

Qy       1  GAGCTATGCTACACACCAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT    60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681  GAGCTATGCTACACACCAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  1740

Qy      61  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGATGGAGAG   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGATGGAGAG  1800

Qy     121  TGAAGTAAACCTGGACCGTTATCAAACAGTTTAGAAGAAGTATTATCGTGGCTTCTTTC    180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801  TGAAGTAAACCTGGACCGTTATCAAACAGTTTAGAAGAAGTATTATCGTGGCTTCTTTC   1860

Qy     181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  1920

Qy     241  CCAGTTTCATACTCATGATGAGGGGTACATGATGGATTGCACGCCCATCAGGGCCGGGTGG   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921  CCAGTTTCATACTCATGATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG  1980

Qy     301  TAATATTCTACAATTGGGAAGTAAGCTGATTGAACAGGAAAAATTATCAGAAGATGAAGA   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981  TAATATTCTACAATTGGGAAGTAAGCTGATTGAACAGGAAAAATTATCAGAAGATGAAGA  2040

Qy     361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  2100

Qy     421  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT    480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT   2160

Qy     481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGAGGAGGAA   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGAGGAGGAA  2220

Qy     541  GCCTCTTGGACCTGATCTTGAAGACTTAAAACGCCAAGTACAACAACATAAGGTGCTTCA   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221  GCCTCTTGGACCTGATCTTGAAGACTTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  2280

Qy     601  AGAAGATCTAGAACAAGAACAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281  AGAAGATCTAGAACAAGAACAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT  2340

Qy     661  TGATGAATCTAGTGGAGATCACGACATCTGAGATGGACAGAGACCCGCTGGGTTCTTTTACAAGACAT   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341  TGATGAATCTAGTGGAGATCACGACATCTGAGATGGACAGAGACCCGCTGGGTTCTTTTACAAGACCA  2400

Qy     721  AGATCGATGGGCAAATCATCTGACTCGTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA   780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401  AGATCGATGGGCAAATCATCTGACTCGTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  2460

Qy     781  CCTTCTCCAAATGGCAACGTCTTACTGACAAAACAGTTCACAACAACTGGCTTTAAAGATCAAAATGAAATGTT   840
            ||||||||||||||||||| |||||||||||||  ||||||||||
Db    2461  GCCTGAC--------------CTAGCTCGTCCTGACTGACCAACTGGCTTTAAAGATCAAAATGAAATGTT  2508

Qy     841  AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT   900
                                        ||||||||||  ||||||||||
Db    2509  GACTGTTACTCTGGTGACACA-------------ACCTGTGGTTACTCAAGGAACATGCCATCT-  2558

Qy     901  ATCCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAGAAGAAAGCAATCCAT   960
            ||||  |||||
Db    2559  ---------------------------------------------------------------  2558
```

```
Qy    1321  CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTCGTTTACAAGACGTTT  1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2794  CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTCGTTTACAAGACGTTT  2853

Qy    1381  GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA  1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2854  GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA  2913

Qy    1441  TTTGGAAGCCAGTTGCACCAGTGGAAGCGGTCTGACCCTTCTCTGCAGGAACTTCTGT  1500
            ||||||||||| || |||||||||||||||||||||||||||||||||||||||||||||
Db    2914  TTTGGAAGCCAGTTGCACCAGTGGAAGCGGTCTGACCCTTCTCTGCAGGAACTTCTGT  2973

Qy    1501  GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC  1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2974  GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC  3033

Qy    1561  AGCAGTTCAGAAGCAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA  1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3034  AGCAGTTCAGAAGCAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA  3093

Qy    1621  ACCTGTAATCATGAGTACTCTTGACAGTGTACGAATATTTCTGACAGAGCAGCCTTTGA  1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3094  ACCTGTAATCATGAGTACTCTTGACAGTGTACGAATATTTCTGACAGAGCAGCCTTTGA  3153

Qy    1681  AGGACTAGAGAAACTTCTACCAGGAGGAGCTGCCTCCTGAGGAGGAGAGCCCCAGAA  1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3154  AGGACTAGAGAAACTTCTACCAGGAGGAGCTGCCTCCTGAGGAGGAGAGCCCCAGAA  3213

Qy    1741  TGTCACTCGGCTTCTACGAAAGCAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA  1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3214  TGTCACTCGGCTTCTACGAAAGCAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA  3273

Qy    1801  CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT  1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3274  CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT  3333

Qy    1861  TCAAGAGGGCCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC  1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3334  TCAAGAGGGCCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC  3393

Qy    1921  CTGGCAGCCGCCGTGGGGCATCTCCTCATTGACTCCTCCAAGATCACCTCGAGAAGTCAA  1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3394  CTGGCAGCCGCCGTGGGGCATCTCCTCATTGACTCCTCCAAGATCACCTCGAGAAGTCAA  3453

Qy    1981  GGCACTTGAGGAGAAATTGCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC  2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3454  GGCACTTGAGGAGAAATTGCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC  3513

Qy    2041  TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGA  2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3514  TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGA  3573

Qy    2101  C  2101
            |
Db    3574  C  3574
```

RESULT 5
JS-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1

```
Qy   2041  TCGCCAGGTTACCACTTTGGGCATTCAGCTTCCACCGTATAACCTCAGCACTCTGAAGA  2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3538  TCGCCAGGTTACCACTTTGGGCATTCAGCTTCCACCGTATAACCTCAGCACTCTGAAGA  3597

Qy   2101  C  2101
           |
Db   3598  C  3598
```

RESULT 6
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9

  Query Match   68.7%; Score 1443; DB 10; Length 3858;
  Best Local Similarity 84.6%; Pred. No. 0;
  Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCTACACGGAGCCCATT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCTACACGGAGCCCATT   959

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGAGTTCATTGATGAGGAG   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGAGTTCATTGATGAGGAG 1019

Qy   121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 1079

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGGTGAAGTCGTGAAGA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGGTGAAGTCGTGAAAGA 1139

Qy   241 CCAGTTCATACTCATGAGGGGTACATGATGGGTACATGATGGGTACATGATGGGTTGG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTCATACTCATGAGGGGTACATGATGGGTACATGATGGGTACATGATGGGTTGG 1199

Qy   301 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy   421 TAGGAGAGTTGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 TAGCAGAGTTGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1379

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAAGAACAAGGAGAAATGAGGAAGA   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAAGAACAAGGAGAAATGAGGAAGA 1439

Qy   541 GCCTCTTGGACCTGATCTTGAAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 GCCTCTTGGACCTGATCTTGAAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1499
```

And the left-side alignment block:

```
Qy    961  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  1020
                       |||   ||  |||| ||   ||
Db   2559  ------------------CCAAACTAGAAAATGCCATCTTCCTGAGTGTGGAG--------  2593

Qy   1021  CCAGAAGACGGAAGCATGGCTGATAACTTTGCCCGGTGTGTTGGGATAATTTAGTCCAAAA  1080
Db   2594  ----------------------------------------------------------  2593

Qy   1081  ACTTGAAAAGAGTACAGCACAGACTCATCAGATTACTTGCAACAGTTCCCCTGGACCTGGA  1140
                ||   ||||||||||||||||||||||||||||||||||||
Db   2594  -------GTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA  2637

Qy   1141  AAAGTTTCTTGCCTGGCTTACAGAGACTGAAACACTGCCAATGTCCTACAGGATGCTAC  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2638  AAAGTTTCTTGCCTGGCTTACAGAGACTGAAACAACTGCCAATGTCCTACAGGATGCTAC  2697

Qy   1201  CCGTAAGGAAAAGGTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA  1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2698  CCGTAAGGAAAAGGTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA  2757

Qy   1261  AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2758  AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG  2817

Qy   1321  CCAAAAAATCCTGAGATCCCTGAAGGTTCCGATGATGCAGTCCTGTTACAAAGAGACGTTT  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2818  CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGAGACGTTT  2877

Qy   1381  GGATAACATGAACTTCAAGTGTGAGTGAACTTCGAGGAACTTCTCTCAACAATTAGGTCCCA  1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2878  GGATAACATGAACTTCAAGTGTGAGTGAACTTCGAGGAACTTCTCTCAACAATTAGGTCCCA  2937

Qy   1441  TTTGGAAGCCAGTTCGACCAGTGGAAGGTGGAAGGTGAAGGTCTCGAGGAACTTCTGGT  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2938  TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGACCAGGAACTTCTCTGAGGAACTTCTGGT  2997

Qy   1501  GTGGCTACAGCTGAAAGATGATGAAATTAAGCACGGCACCTATTGGAGGCGACTTTCC  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2998  GTGGCTACAGCTGAAAGATGATGAAATTAAGCACGGCACCTATTGGAGGCGACTTTCC  3057

Qy   1561  AGCAGTTCAGAAGCAGAAGACGAACAGATGTACATAGGGGGCTTCAAGGGGAATTGAAAACTAAAGA  1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3058  AGCAGTTCAGAAGCAGAAGACGAACAGATGTACATAGGGGGCTTCAAGGGGAATTGAAAACTAAAGA  3117

Qy   1621  ACCTGTAATCATGAGTACTCTTGAGACTCGTACGAATATTTCTGACAGAGCCAGCCTTTGGA  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3118  ACCTGTAATCATGAGTACTCTTGAGACTCGTACGAATATTTCTGACAGAGCCAGCCTTTGGA  3177

Qy   1681  AGGACTAGAGAAACTCTACCAGGAGCCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAA  1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3178  AGGACTAGAGAAACTCTACCAGGAGCCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAA  3237

Qy   1741  TGTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAATTGAA  1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3238  TGTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA  3297

Qy   1801  CCTGCACTCCGCTGACTGGCAGAGAAAATGAGACCCTTGAAAGACTTCAGGAACT  1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3298  CCTGCACTCCGCTGACTGGCAGAGAAAATAGATGGGAATGCCTCAGGGTAGC  3357

Qy   1861  TCAAGAGGGCCACGGATGAGCTGAGCTGGACCTCCAAGCTGCGCGCAAGCTGAGGTGATCAAGGGATC  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3358  TCAAGAGGGCCACGGATGAGCTGAGCTGGACCTCCAAGCTGCGCGCAAGCTGAGGTGATCAAGGGATC  3417

Qy   1921  CTGGCAGCCCCTGGGCGCGGGCGATCTCCATTGACTCCTCCAAGATCACCTCGAGAAAGTCAA  1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3418  CTGGCAGCCCCTGGGCGCGGGCGATCTCCATTGACTCCTCCAAGATCACCTCGAGAAAGTCAA  3477

Qy   1981  GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAAGACGTCAATGACCTTGC  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3478  GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAAGACGTCAATGACCTTGC  3537
```

Qy 1681 AGGACTAGAGAAAACTTCTACCAGGAGCTGCCTCCTGAGAGAGCGCCCAGAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2256 AGGACTAGAGAAAACTTCTACCAGGAGCTGCCTCCTGAGAGAGCGCCCAGAA 2315

Qy 1741 TGTCACTCGGCTTCACGAAGAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2316 TGTCACTCGGCTTCACGAAGAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 2375

Qy 1801 CCTGACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2376 CCTGACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 2435

Qy 1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2436 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 2495

Qy 1921 CTGGCAGCCGTGGCGACTCTCTCATTGACTCTTCCAAGATCACCTCGAGAAAGTCAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2496 CTGGCAGCCGTGGCGACTCTCTCATTGACTCTTCCAAGATCACCTCGAGAAAGTCAA 2555

Qy 1981 GGCACTTGCGAGGAGAAATTGCGCCTCTGAAAAGAGAAACGTGAGCCCAGTCAATGACCTTGC 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2556 GGCACTTGCGAGGAGAAATTGCGCCTCTGAAAAGAGAAACGTGAGCCCAGTCAATGACCTTGC 2615

Qy 2041 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACGTATAACCTCAGCACTCTGGAGA 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2616 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACGTATAACCTCAGCACTCTGGAGA 2675

Qy 2101 C 2101
Db 2676 C 2676

RESULT 7
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match   68.7%; Score 1443; DB 10; Length 4825;
Best Local Similarity   84.6%;
Matches 1777; Conservative   0; Mismatches   0; Indels 324; Gaps 1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1716

Qy   61 TCCTTCACAGCATTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1717 TCCTTCACAGCATTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1776

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTCAAAACAGCTTTAGAAGAAGAGATTTCTAATGATGTGGAGAGTGGTGAAAAGA   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTCAAAACAGCTTTAGAAGAAGAGATT 1836

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGACACACAAGGAGATTCTAATGATGTGGAGAGTGGTGAAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1837 TGCTGAGGACACATTGCAAGCACAAGGACACACAAGGAGAGATTTCTAATGATGTGGAGAGTGGTGAAAGA 1896

Left column (RESULT 6 continuation):

Qy   601 AGAAGATCTAGAACAAGAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGGTAGT   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1500 AGAAGATCTAGAACAAGAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGGTAGT 1559

Qy   661 TGATGAATCAGTGGAATCACGGACAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1560 TGATGAATCAGTGGAATCACGGACAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAT   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1620 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGAC-- 1677

Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1677 --- 1677

Qy   841 AAAAGAGAGATCCAGTGAACAGAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1678 --- 1678

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGACCGATCTAGAAAGAAAGCAATCCAT   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1678 --- 1678

Qy   961 GGGCAAACTGTATTCACTCAAACAGATATCTTCTTCACACTGAAGAATAAGTCAGTGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1678 -- 1678

Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTGTCCAAAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1678 -- 1678

Qy 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGACCTTGGA 1140
          ||||||||||||||||||||||||| ||||||||||||||||
Db 1678 -------------------------ACTCATAGATTACTGCAACAGTTCCCCTGACCTTGGA 1715

Qy 1141 AAAGTTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGCTCCTACAGGATGTGTAC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1716 AAAGTTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGCTCCTACAGGATGTGTAC 1775

Qy 1201 CCGTAAGGCAAAGGCTCTAGAAGAGTTCCAAGGAGTAAAAGAGCTGATGAAACAATGCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1776 CCGTAAGGCAAAGGCTCTAGAAGAGTTCCAAGGAGTAAAAGAGCTGATGAAACAATGCA 1835

Qy 1261 AGACCTTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1836 AGACCTTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAG 1895

Qy 1321 CCAAAAAAATCTGAGATCCCTGGAGGTTCCGATGAGTCCAGTCCTGTTACAAAGACGTTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1896 CCAAAAAAATCTGAGATCCCTGGAGGTTCCGATGAGTCCAGTCCTGTTACAAAGACGTTT 1955

Qy 1381 GGATAACATGATGAGCTTCAAGTGGAGTGGAGTGAACTTCGGAAAAAAGTCTCTCAACATTAGGTCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1956 GGATAACATGATGAGCTTCAAGTGGAGTGGAGTGAACTTCGGAAAAAAGTCTCTCAACATTAGGTCCCA 2015

Qy 1441 TTTGGAAGCCAGTTCTGACCAGTCGAGGCGTCTGCAGGAACTTCTGGT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||
Db 2016 TTTGGAAGCCAGTTCTGACCAGTCGAGGCGTCTGCAGGAACTTCTGGT 2075

Qy 1501 GTGGCTACAGCTGAAAGATGATGAAGCTCATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2076 GTGGCTACAGCTGAAAGATGATGAAGCTCATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 2135

Qy 1561 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2136 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 2195

Qy 1621 ACCTGTAATCAATGAGTACTCTTGAGACTGTGACGAATATTTCTGACAGAGCAGCCTTCGGA 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2196 ACCTGTAATCAATGAGTACTCTTGAGACTGTGACGAATATTTCTGACAGAGCAGCCTTCGGA 2255

```
Qy    241 CCAGTTTCATACTCCATGAGGGGTACATGATGGATTTGACAGCCATCAGGGCCGGGGTTGG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTTCATACTCCATGAGGGGTACATGATGGATTTGACAGCCATCAGGGCCGGGGTTGG 1956

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA 2016

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATCCTCAGGGTAGC  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATCCTCAGGGTAGC 2076

Qy    421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAAACT  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAAACT 2136

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAGAACAGGAGAAAATGAGGAAGA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAGAACAGGAGAAAATGAGGAAGA 2196

Qy    541 GCCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATCACATGGTGGTAGT  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATCACATGGTGGTAGT 2316

Qy    661 TGATGAATCTAGTGCAATCACGGACAATCTGCTCTTGAAGACAACTAAGGTATTGGG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TGATGAATCTAGTGCAATCACGGACAATCTGCTCTTGAAGACAACTAAGGTATTGGG 2376

Qy    721 AGATCGATGGGCAAACATCGTGTAGATGGACAGAAGACCGCTGGGTTCCTTTTACAGAGACAT  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 AGATCGATGGGCAAACATCGTGTAGATGGACAGAAGACCGCTGGGTTCCTTTTACAGAGAC-- 2434

Qy    781 CCTTCTCAAATGGCAACGTCTTACTGAAGAAGTAGCAGTGCCTTTTGTTAGTGCATGGCTTTCAGA  840
          
Db   2434 ------------------------------------------------------------------ 2434

Qy    841 AAAAGAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAAATGTT  900
          
Db   2434 ------------------------------------------------------------ 2434

Qy    901 ATCAAGTCTTCAAAAACTGGCCGGTTTAAAGCGGATCTAGAAAAAGAAAAAAGCAATCCAT  960
          
Db   2434 ------------------------------------------------------------ 2434

Qy    961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
          
Db   2434 ------------------------------------------------------------ 2434

Qy   1021 CCAGAAACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
          
Db   2434 ------------------------------------------------------------ 2434

Qy   1081 ACTTGAAAAGAGTACAGCAGTACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGA 1140
          |||||||||||||||||||||||||||||||||||
Db   2435 ------------------ACTTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2472

Qy   1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACACAGGATGCTAC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2473 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCCAATGTCCTACACAGGATGCTAC 2532

Qy   1201 CCGTAAGGAGAAGGCTCCTAGAGAACTCCAGGGAGTAAAAGAGCTGATGAAACAATGGCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2533 CCGTAAGGAGAAGGCTCCTAGAGAACTCCAGGGAGTAAAAGAGCTGATGAAACAATGGCA 2592

Qy   1261 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2593 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG 2652

Qy   1321 CCAAAAAATCCTGAGATCCTGGAAGGTTCGGATGATGCAGTCCTGTTACAAAGACGTTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2653 CCAAAAAATCCTGAGATCCTGGAAGGTTCGGATGATGCAGTCCTGTTACAAAGACGTTT 2712

Qy   1381 GGATAACACATGAACTTCAAGTGAGTGAACTTCGGAGAACTTCCTCAACATTAGGTCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2713 GGATAACACATGAACTTCAAGTGAGTGAACTTCGGAGAACTTCTCAACATTAGGTCCCA 2772

Qy   1441 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCCACGTTGTCCTGCGAGGAACTTCTTGGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2773 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCCACGTTTCTCTGCGAGGAACTTCTTGGT 2832

Qy   1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGGCGCACTTTCC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2833 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGGCGCACTTTCC 2892

Qy   1561 AGCAGTTCAGAAGCAGAACGATGTACATGAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2893 AGCAGTTCAGAAGCAGAACGATGTACATGAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 2952

Qy   1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2953 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 3012

Qy   1681 AGGACTAGAGAAACTCTACCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3013 AGGACTAGAGAAACTCTACCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAA 3072

Qy   1741 TGTCACTCGGCTTCACGAGAAAGCAGGCTGAGGAGTCAATACTGAGTGTGGGAAAATTGAA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3073 TGTCACTCGGCTTCACGAGAAAGCAGGCTGAGGAGTCAATACTGAGTGTGGGAAAATTGAA 3132

Qy   1801 CCTGCACTCCGCTGACTGGCAGAGAAAATGAGACCCTTGAAAGACTCCAGGAACT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3133 CCTGCACTCCGCTGACTGGCAGAGAAAATGAGACCCTTGAAAGACTCCAGGAACT 3192

Qy   1861 TCAAGAGGGCACGGATGAGCTGACCTGCCAAGTCTGCGCCAAGTGAGGTGATCAAGGGATC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3193 TCAAGAGGGCACGGATGAGCTGACCTGCCAAGTCTGCGCCAAGTGAGGTGATCAAGGGATC 3252

Qy   1921 CTGGCAGCCCGTGGGCGATCTCCTCATTGCAGTTCTCGAGAAGAACGCTGAGCCACGTCAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3253 CTGGCAGCCCGTGGGCGATCTCCTCATTGCAGTTCTCGAGAAGAACGCTGAGCCACGTCAA 3312

Qy   1981 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3313 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3372

Qy   2041 TCGGCCAGCTTACCACTTGGGCATTCAGCTCTCCACCGTATAACCTCACTCTGGAAGA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3373 TCGGCCAGCTTACCACTTGGGCATTCAGCTCTCCACCGTATAACCTCACTCTGGAAGA 3432

Qy   2101 C 2101

Db   3433 C 3433
```

```
LENGTH: 4848
TYPE: DNA
ORGANISM: Homo sapiens
S-09-845-416-35

Query Match        68.7%;   Score 1443;  DB 10;  Length 4848;
Best Local Similarity  84.6%;   Pred. No. 0;
Matches 1777;   Conservative  0;   Mismatches  0;   Indels 324;   Gaps  1;

Qy    1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCGACCCACACGGAGCCCATT   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1680 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCGACCCACACGGAGCCCATT 1739

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1740 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1799

Qy  121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1800 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1859

Qy  181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1860 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1919

Qy  241 CCAGTTCATATCTCATGAGGGGTACAGCAGGGATTGACAGCCCATCAGGGCCGGGTTGG  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1920 CCAGTTCATATCTCATGAGGGGTACAGCAGGGATTGACAGCCCATCAGGGCCGGGTTGG 1979

Qy  301 TAATATTCTACAATTGGGAAGTAAGCTAAGCAGCTTAGAAAATTATCAGAAGAGATGAAGA  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1980 TAATATTCTACAATTGGGAAGTAAGCTAAGCAGCTTAGAAAATTATCAGAAGAGATGAAGA 2039

Qy  361 AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2040 AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2099

Qy  421 TAGCATGGAAAAGCAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2100 TAGCATGGAAAAGCAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2159

Qy  481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAGAAATGGAGGAAGA  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2160 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAGAAATGGAGGAAGA 2219

Qy  541 GCCTCCTTGAACCTGATCTTGAAGACTTACCAGGTACAAACGCCAAGTACAACAACATAAGGTGCTTCA  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2220 GCCTCCTTGAACCTGATCTTGAAGACTTACCAGGTACAAACGCCAAGTACAACAACATAAGGTGCTTCA 2279

Qy  601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTTACAACGGTGTGGTAGT  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2280 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTTACAACGGTGTGGTAGT 2339

Qy  661 TGATGAATCTAGTGAAGACTTAGTGAGATCACGCAACTGTGCTTGCTTGGAAGACAACAACTTAAGGTATTGGG  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2340 TGATGAATCTAGTGAAGACTTAGTGAGATCACGCAACTGTGCTTGCTTGGAAGACAACAACTTAAGGTATTGGG 2399

Qy  721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGACAT  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2400 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGAC-- 2457

Qy  781 CCTTCCCAAATGGCAACGTCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  840
        --------------------------------------------------------------
Db 2458 -------------------------------------------------------------- 2457

Qy  841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT  900
        ------------------------------------------------------------
Db 2458 ------------------------------------------------------------ 2457

Qy  901 ATCAAGTACTCTCAAAAACTGGCGGTTTAAAAAGCCGGATCTAGAAAGGAAAAAGCAATCCAT  960
        ------------------------------------------------------------
Db 2458 ------------------------------------------------------------ 2457

Qy  961 GGGCAAACTGTATTCACTCACAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
        ------------------------------------------------------------
Db 2458 ------------------------------------------------------------ 2457

Qy 1021 CCAGAGACGGAAGCATGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
        --------------------------------------------------------
Db 2458 -------------------------------------------------------- 2457

Qy 1081 ACTTGAAAAGAGTACAGCAGACTCATAGATTACTGCAACAGTTCCCCTGGACTGGA 1140
        ------|||||||||||||||||||||||||||||||||||||
Db 2458 ------ATCTATAGATTACTGAACAGTTCCCCTGGACCTGGA 2495

Qy 1141 AAAGTTCTTGCCTGGCTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2496 AAAGTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 2555

Qy 1201 CCGTAAGCAAAGGCTCCAAGGAGACTCCAAGGAGTAAAGAGCTGATGAAACAATGGCA 1260
        ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db 2556 CCGTAAGGAAAGGCTCCAAGGAGACTCCAAGGAGTAAAGAGCTGATGAAACAATGGCA 2615

Qy 1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACAACTGGATGAAAACAG 1320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2616 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACAACTGGATGAAAACAG 2675

Qy 1321 CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGAGTCAGTCGTTGTTACAAAGACGTTT 1380
        |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db 2676 CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGAGTCAGTCGTCCTGTTACAAAGACGTTT 2735

Qy 1381 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGGTCCCA 1440
        ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db 2736 GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGGTCCCA 2795

Qy 1441 TTTGAAGCCAGTTCGACCAGTGAGCGTCTGCGAGCGTCTGCACCTTCTCTGGAACTTCTGGT 1500
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2796 TTTGAAGCCAGTTCGACCAGTGAGCGTCTGCGAGCGTCTGCACCTTCTCTGCAGGACTTCTGGT 2855

Qy 1501 GTGGCTACAGCTGAAAGATGATGAAAGATGAATTAAGCCCGGCAGGCACCTATTGGAGCGCGACTTTCC 1560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2856 GTGGCTACAGCTGAAAGATGATGAAAGATGAATTAAGCCCGGCAGGCACCTATTGGAGCGCGACTTTCC 2915

Qy 1561 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGAAAATGAGAGGAAGA 1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2916 AGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGAAAATGAGAGGAAGA 2975

Qy 1621 ACCTGTAATCATGAGTACTCTTGAGACTGAACATAGTCTTCGACAATATTTCTGGA 1680
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2976 ACCTGTAATCATGAGTACTCTTGAGACTGTACATAGTCTTCGACAGCAGCCTTTGGA 3035

Qy 1681 AGGAACTAGAATAAACGCTACCAGAGAGCTGCCTCCTGGAGAGAGCCCAGAA 1740
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3036 AGGACTAGAGAAAC-CTACCAGAGAGCTGCCTCCTGAGTGGAGCGCCCAGAA 3095

Qy 1741 TGTCACTCGGCTTCTCACTGCCAACTTAGGTACAATATTTGGA 1800
        ||||||||||||||||||||||||||||||||||||||||||||
Db 3096 TGTCACTCGGCTTCTCACTGCCAATACCGGTACAATATTTGAA 3155

Qy 1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGAGACCCTTGAAGACTCCAGGAACT 1860
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3156 CCTGCACTCCGCTGACTGGCAGAGAAATAGAGACCCTTGAAGACTCCAGGAACT 3215

Qy 1861 TCAAGAGGGCACCGATGAGCTGGACCTCAAGTGATCAAGGGATC 1920
        ||||||||||||||||||||||||||||||||||||||||||||
Db 3216 TCAAGAGGGCACCGATGAGCTGGACCTCAAGTGATCAAGGGATC 3275

Qy 1921 CTGGGCACCCGTGGCGATCTCCTCATTGACTCTCGAGAAAGTTCAA 1980
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 3276 CTGGGCACTCGGCGATCTCCTCATTGACTCTCGAGAAAGTTCAA 3335

Qy 1981 GGCACTTGCGCCTCTGAAAGAGAAGTGAGCCACGTCAATGACCTTGC 2040
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 3336 GGCACTTGAGGAGAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3395

Qy 2041 TCGGCCAGCTTACCACCTTTGGGCATTCAGCTCTCAGCTCTCACCCGTATAACCTCACTCTGGAAGA 2100
```

```
3396  TGCCCAGCTTACCACTTTGGGCATTCAGCTCTCAGCTCTACCGTATAACCTCAGCACTCTGGAAGA  3455
                     |||||||||||||||||||||||||||||||||||||||||||||||||||||||

2101  C 2101
      |
3456  C 3456
```

RESULT 9
US-09-845-416-36
Sequence 36, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
  APPLICANT: XIAO, XIAO
  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
  TITLE OF INVENTION: THEREOF
  FILE REFERENCE: DE1142
  CURRENT APPLICATION NUMBER: US/09/845,416
  CURRENT FILING DATE: 2001-04-30
  PRIOR APPLICATION NUMBER: 60/200,777
  PRIOR FILING DATE: 2000-04-28
  NUMBER OF SEQ ID NOS: 36
  SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
  LENGTH: 5060
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-845-416-36

Query Match   68.7%; Score 1443; DB 10; Length 5060;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

```
Qy     1 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACCGGAGCCCATT    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1892 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACCGGAGCCCATT  1951

Qy    61 TCCTTCACAGCATTGGAAGCTCCTGAGACAGTCATTTGGCAGTCATTTGGCAGTCATTTGGCAGTGAGAG   120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1952 TCCTTCACAGCATTGGAAGCTCCTGAGACAGTCATTTGGCAGTCATTTGGCAGTGATGGAGAG  2011

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGGCTTCTTTC   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2012 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC  2071

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2072 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  2131

Qy   241 CCAGTTCATACTCATGAGGGGTACATGAGGGGTACATGGATTTGACAGCCCATCAGGGGCCGGGTTGG   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2132 CCAGTTCATACTCATGAGGGGTACATGAGGGGTACATGGATTTGACAGCCCATCAGGGGCCGGGTTGG  2191

Qy   301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2192 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA  2251

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2252 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  2311

Qy   421 TAGCATGGAAAAACAAAGCAATTCACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2312 TAGCATGGAAAAACAAAGCAATTCACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  2371

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAACAGAAGAAATTGAGGAAGA   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAACAGAAGAAATTGAGGAAGA  2431

Qy   541 GCCTCTTGGACCTGATCTTGAAGACAATTCAGACCTAAAAACGCCAAGTAACAATAAAGCCTTCA   600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2432 GCCTCTTGGACCTGATCTTGAAGACAATTCAGACCTAAAAACGCCAAGTAACAATAAAGCCTTCA  2491
```

```
Qy   601 AGAGAGATTCAGAACAAGAACAAGACAAGTCAGGGTCAGGGTCAATTCCTCACTCACATGGTGGTGGTAGT   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2492 AGAGAGATTCAGAACAAGAACAAGACAAGTCAGGGTCAATTCCTCACTCACATGGTGGTGGTAGT  2551

Qy   661 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2552 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG  2611

Qy   721 AGATCGATGGGCAAACATCGTAGATGGACAGAAGACCGCTGGGTTCTTTACAAGACAT   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||  --
Db  2612 AGATCGATGGGCAAACATCGTAGATGGACAGAAGACCGCTGGGTTCTTTACAAGAC--  2669

Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCTTTTGTCATGGCTTTCAGA  840
                                                                  ------
Db  2670 ---------------------------------------------------------  2669

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAAATGTT   900
                                                                      ----
Db  2670 -----------------------------------------------------------  2669

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGGCGGATCTAGAAAAGAAAAAAGCAATCCAT   960
                                                                        ----
Db  2670 -----------------------------------------------------------  2669

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCACACTGAAGAATAAGTCAGTGAC  1020
                                                                     ----
Db  2670 ---------------------------------------------------------  2669

Qy  1021 CCAGAAGACGGAAGCATGCTGGATCATGCTGGATAACTTTGCCCGGTGTGTTGGGATAATTTAGTCCAAAA  1080
                                                                          ----
Db  2670 -----------------------------------------------------------  2669

Qy  1081 ACTTGAAAAGAGTACAGACACAGACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA  1140
                           ||||||||||||||||||||||||||||||||||||||||||
Db  2670 --------------ACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA  2707

Qy  1141 AAAGTTTCTTGCCTGGCTTACAGAAGTGAAACAACTGCCAATGTCCTACAGGATGCTAC  1200
                               ||||||||||||||||||||||||||||||||||||||||
Db  2708 --------------AAAGTTTCTTGCCTGGCTTACAGAAGTGAAACAACTGCCAATGTCCTACAGGATGCTAC  2767

Qy  1201 CCGTAAGGAAAGGCTCCTCAGAAGACTCCAAGGGAGTAAAAAGAGCTGATGAAAACAATGGCA  1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2768 CCGTAAGGAAAGGCTCCTCAGAAGACTCCAAGGGAGTAAAAAGAGCTGATGAAAACAATGGCA  2827

Qy  1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAG  1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2828 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAG  2887

Qy  1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2888 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCGTCCTGTTACAAAGACGTTT  2947

Qy  1381 GGATAACATGAACTTCAAGTGACTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA  1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2948 GGATAACATGAACTTCAAGTGACTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA  3007

Qy  1441 TTTGGAAGCCAGTTCTGACCCAGTGGAAGCGTCTGCACCCTTTCCTCGCAGGAACTTCTGGT  1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3008 TTTGGAAGCCAGTTCTGACCCAGTGGAAGCGTCTGCACCCTTTCTCGCAGGAACTTCTGGT  3067

Qy  1501 GTGGCTACAGCTGAAAGATGATGAGAAATGATGAATTAAGCCGGGCAACCTATTGGAGGCGACTTTCC  1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3068 GTGGCTACAGCTGAAAGATGATGAGAAATGATGAATTAAGCCGGGCAACCTATTGGAGGCGACTTTCC  3127

Qy  1561 AGCAGTTCAGAGGACGAACGATGTACATAGGGGCTTCAAGAGGAATTGAAAACTAAAGA  1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3128 AGCAGTTCAGAGGACGAACGATGTACATAGGGGCTTCAAGAGGAATTGAAAACTAAAGA  3187

Qy  1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA  1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3188 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA  3247

Qy  1681 AGGACTAGAGAAACTCTACCAGAGACCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAA  1740
```

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGACAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGACAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATACTCATGAGGGTACATGATGAATTGACAGCCCATCAGGGCCGGGTTGG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTCATACTCATGAGGGTACATGATGAATTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGAACACAGAAAATTATCAGAGATGAAGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGAACACAGAAAATTATCAGAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA 1638

Qy   541 GCCTCTTGGACCTGATCTTGAAGACTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1698

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGTGGTAGT 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGTGGTAGT 1758

Qy   661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTAAGGTATTGGG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTAAGGTATTGGG 1818

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGAACCGCTGGGTTCTTTACAAGACAT 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGAACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   781 CCTTCTCAAATGGCAAGTCTTACTGAAGACAGTGCCTTTTAGTGCATGGCTTCAGA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAAGTCTTACTGAAGACAGTGCCTTTTTAGTGCATGGCTTCAGA 1938

Qy   841 AAAAGAAGATGCAGTGAACAAGATCCACAACTGGCTTAAGATCAAAATGAAATGTT 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAAGATCCACAACTGGCTTAAGATCAAAATGAAATGTT 1998

Qy   901 ATCAAGTCTTCAAAAACTGCCGTTTAAAAGGGGATCCAGAAAAAGAAAAAGCAATCCAT 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGCCGTTTAAAAGGGGATCCAGAAAAAGAAAAAGCAATCCAT 2058

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 2118

Qy   1021 CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGTTGGGATAATTTAGTCCAAAA 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGTTGGGATAATTTAGTCCAAAA 2178

Qy   1081 ACTTGAAAAGAGTACAGCACAG------ 1102
           ||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGATTTCACAGCAGCGCTCCTGACTGACCAC 2238

Qy   1103 ------ 1102
Db   2239 TATTGGAGCCTCTCCTACTCAGACTGTTACTCGGTGACAGAACCTGGTTACTAAGGA 2298

Qy   1103 ------ 1102
Db   2299 AACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGAGTGTTGGAGGTACCTGCTCTGGC 2358

Qy   1103 ------ 1102

b   3248 AGGACTAGAGAAAACTCTACCAGGAGACGACCCCAGAGAGCTGCCTCCTCTGAGGAGAGAGCCCAGAA 3307
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
y   1741 TGTCACTCGGCTTCTACGAAAGCAAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3308 TGTCACTCGGCTTCTACGAAAGCAAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3367

y   1801 CCTGCACTCCCGTGACTGGCAGAGAGAAAATAGATGAGACTCCAGAACT 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||
b   3368 CCTGCACTCCCGTGACTGGCAGAGAGAAAATAGATGAGACTCCAGAACT 3427

y   1861 TCAAGAGGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3428 TCAAGAGGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 3487

y   1921 CTGGCAGCGCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACTCGAGAAGTCAA 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3488 CTGGCAGCCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACTCGAGAAGTCAA 3547

y   1981 GGCACTTCGAGGAGAAATTGCGCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTGC 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3548 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3607

y   2041 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3608 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3667

y   2101 C 2101
b   3668 C 3668

ESULT 10
S-10-149-736-42
Sequence 42, Application US/10149736
Publication No. US2003021632A1
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR APPLICATION NUMBER: US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: Patentin version 3.1
SEQ ID NO 42
LENGTH: 8689
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
S-10-149-736-42

Query Match   61.5%;   Score 1293; DB 15;   Length 8689;
Best Local Similarity   72.5%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 798; Gaps 1;

y   1 GAGCTATGCCTACACACAGCGCTGCTTATGTCACCACCTTGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1099 GAGCTATGCCTACACACAGCGCTGCTTATGTCACCACCTTGACCCTACACGGAGCCCATT 1158

y   61 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1159 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1218

y   121 TGAAGTAAACCTGGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1219 TGAAGTAAACCTGGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

```
Db  2359  AGATTTCAACCGGGCTTGGACAGAACTTACCGACTGGCTTTCTCTGCTGCTGATCAAGTTAT  2418
Qy  1103  -------------------------------------------------------------  1102

Db  2419  AAAATCACAGAGGGTGATGGTGGGTGACCTTGAGGATATCAACGAGATGATGATCATCAAGCA  2478
Qy  1103  -------------------------------------------------------------  1102

Db  2479  GAAGGCAACAATGCAGGATTTGGAAACAGAGAGCGTCCCCAGTTGGAAGAACTCATTACCGC  2538
Qy  1103  -------------------------------------------------------------  1102

Db  2539  TGCCCAAAATTTGAAAAACAAGACCAGCAATCAAGAGGCTAGAACAATCATTACCGGATCG  2598
Qy  1103  -------------------------------------------------------------  1102

Db  2599  AATTGAAAGAATTCAGAATCAGTGGAGTACAAGAACACCTTCAGAACCGGAGGCA  2658
Qy  1103  -------------------------------------------------------------  1102

Db  2659  ACAGTTGAATGAAATGTTAAAGGATTCAAACACAATGGCTGGAAGCTAAGGAAGAAGCTGA  2718
Qy  1103  -------------------------------------------------------------  1102

Db  2719  GCAGGTCTTAGGACAGGCCAGAGCCAAGCTTGAGTCATGATGGAGGAGGGTCCCTATACAGT  2778
Qy  1103  -------------------------------------------------------------  1102

Db  2779  AGATGCAATCCAAAAGAAAAATCACAGAAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTG  2838
Qy  1103  -------------------------------------------------------------  1102

Db  2839  GCGAGACAAAATGTAGATGTGGCAAAATGACTTGGCCCTGAAATGAACTTCTCTCGGGATTATTCTGC  2898
Qy  1103  -------------------------------------------------------------  1102

Db  2899  AGATGATACCAGAGAAAGTCCACATGAGAATATCAATGCGCTCTTGGAGAAGCAT  2958
Qy  1103  -------------ACTCATAGATTACTGCAACA  1122

Db  2959  TCATAAAGGGGTGAGTGAGCGAGAGGCTGCTTTGCCTGGCTTGCCTGGCTGAAACAACTGCCAA  3018
Qy  1123  GTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTGCCTGGCTGAAACAACTGCCAA  3078

Db  1123  GTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTGCCTGGCTGAAACAACTGCCAA  1182
Qy  3019  GTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTGCCTGGCTGAAACAACTGCCAA  3078

Db  1183  TGTCCTACAGGATGCGTACCCCGTAAGGAAAAGGCTCTAAGAGACTCCAAGGAGTAAAAGA  1242
Qy  3079  TGTCCTACAGGATGCGTACCCCGTAAGGAAAAGGCTCTAAGAGACTCCAAGGAGTAAAAGA  3138

Db  1243  GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCA  1302
Qy  3139  GCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCA  3198

Db  1303  CAACCTGTTACAAAGACGTTGTTTGGATGATAACATGAACATGAACATGACTTCAAGTGAAGTC  1362
Qy  3199  CAACCTGTTACAAAGACGTTGTTTGGATGATAACATGAACATGAACATGACTTCAAGTGAAGTC  3258

Db  1363  CCTGTTACAAAGACGTTTGGATGATAACATGAACATGAACATGGATCCCTGGAAGTGAACTTCAAGTGAGTGAGTGAACTTCGGAAAAAGTC  1422
Qy  3259  CCTGTTACAAAGACGTTTGGATGATAACATGAACATGAACATGGATCCCTGGAAGTGAACTTCAAGTGAGTGAGTGAACTTCGGAAAAAGTC  3318

Db  1423  TCTCAACATTAGGTCCCATTGGAAGCCAGTCTGACCAGTGAAGCGTCTGCACCTTTC  1482
Qy  3319  TCTCAACATTAGGTCCCATTGGAAGCCAGTCTGACCAGTGAAGCGTCTGCACCTTTC  3378

Db  1483  TCTGCAGGAACTTCTCTGGTGTGCTACAGCTGCAACTGATGAAAAGATGATTGAATTAAGCCGGCAGGCACC  1542
Qy  3379  TCTGCAGGAACTTCTCTGGTGTGCTACAGCTGCAACTGATGAAAGATGATTGAATTAAGCCGGCAGGCACC  3438

Db  1543  TATTGGAGGCGACTTTCCAGCAGTTCAGAAGCTGTACATAGGGCCTTCAAGAG  1602
```

RESULT 11
US-09-845-416-32
```
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
;  APPLICANT: XIAO, XIAO
;  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: DE1142
;  CURRENT APPLICATION NUMBER: US/09/845,416
;  CURRENT FILING DATE: 2001-04-30
;  PRIOR APPLICATION NUMBER: 60/200,777
;  PRIOR FILING DATE: 2000-04-28
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
;  LENGTH: 4414
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-845-416-32

Query Match     53.0%;  Score 1114;  DB 10;  Length 4414;
Best Local Similarity   97.0%;
Matches 1135;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;
```

```
                                                                        Pred. No. 2.9e-313;

Qy     1  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1657  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1716

Qy    61  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTGCAGTTCATTGATGGAGAG       120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1717  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTGCAGTTCATTGATGGAGAG      1776

Qy   121  TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC   180
```

(Right column alignment blocks)

```
Db  3439  TATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGACGATGTACATAGGGCCTTCAAGAG  3498
Qy  1603  GGAATTGAAAAACTAAAGAACCTGTAATCAATCATGAGTACTCTTGAGACTGTACGAATATTTCT  1662
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3499  GGAATTGAAAAACTAAAGAACCTGTAATCAATCATGAGTACTCTTGAGACTGTACGAATATTTCT  3558

Qy  1663  GACAGAGCAGCCTTTGGAAGACTAGAGAACTCTACCAGGAGCCCAGAGAGCTGCCTCC  1722
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3559  GACAGAGCAGCCTTTGGAAGACTAGAGAACTCTACCAGGAGCCCAGAGAGCTGCCTCC  3618

Qy  1723  TGAGGAGAGAGCCCAGAATGTCACTGGCTTCACGAAAAGCAGGGTGAGGAGGTCAATAC  1782
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3619  TGAGGAGAGAGCCCAGAATGTCACTCGGCTTCACGAAAAGCAGGGTGAGGAGGTCAATAC  3678

Qy  1783  TGAGTGGGAAAAATTGAACCTGCTACTCGCTGACTCGGCAGAGAAAAATAGATGAGACCCT  1842
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3679  TGAGTGGGAAAAATTGAACCTGCTACTCGCTGACTCGGCAGAGAAAAATAGATGAGACCCT  3738

Qy  1843  TGAAAGACTTCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTTCAAGCTGCGCCAAGC  1902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3739  TGAAAGACTTCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTTCAAGCTGCGCCAAGC  3798

Qy  1903  TGAGGTGATCAAGGATCCTGGCAGCCGTGGGCGTCCATCTCCTTCATTGACTTCCTCAAGA  1962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3799  TGAGGTGATCAAGGATCCTGGCAGCCGTGGGCGTCCATCTCCTTCATTGACTTCCTCAAGA  3858

Qy  1963  TCACCTCCAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACCTGAG  2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3859  TCACCTCCAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACCTGAG  3918

Qy  2023  CCACGTCAATGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAA  2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3919  CCACGTCAATGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAA  3978

Qy  2083  CCTCAGACTCTGGAAGAC  2101
          ||||||||||||||||||
Db  3979  CCTCAGACTCTGGAAGAC  3997
```

Query Match 52.5%; Score 1103.4; DB 15; Length 5417;
Best Local Similarity 99.9%; Pred. No. 4e-310;
Matches 1104; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

```
Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy    61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAGTTCATTGGCAGTTCATTGATGGAGAG  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAGTTCATTGGCAGTTCATTGATGGAGAG 1218

Qy   121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGCTTCTTTC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATATCTCATGAGGGGTACATGATGAGGAAGTAAGCTGATTGAACAGGAAAATTATCAGAAGATGAAGA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTCATATCTCATGAGGGGTACATGATGAGGAAGTAAGCTGATTGAACAGGAAAATTATCAGAAGATGAAGA 1398

Qy   301 TAATATTCTACAATTGGAAGTAAGCTGATTGAACAGGAAAATTATCAGAAGATGAAGA  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1399 TAATATTCTACAATTGGAAGTAAGCTGATTGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459 AACTGAAGTACAAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAACAAAGCAATTTACTAGAGTTTAATGAGTTTAATGGATCTCCAGAATCAGAAACT  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1519 TAGCATGGAAAAACAAAGCAATTTACTAGAGTTTAATGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   481 GAAAGAGTTGAATGACTGGCTAACAACAGAAGAAGAACAGAAGAAAAATGGAGAAGA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579 GAAAGAGTTGAATGACTGGCTAACAACAGAAGAAGAACAGAAGAAAAATGGAGAAGA 1638

Qy   541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGACCAAGTACAACAACATAAGGTGCTTCA  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGACCAAGTACAACAACATAAGGTGCTTCA 1698

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCAATTCTCCACATGGTGGTAGT  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCAATTCTCCACATGGTGGTAGT 1758

Qy   661 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTTGGAAGAACAACTTAAGGTATTGGG  720
```

```
b    1777 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGCTTCTTTC 1836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
y     181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

y     241 CCAGTTCATATCTCATGAGGGGTACATGATGAGGAAGTAAGCTGATTGGAAGTGCCTCAGGGGTTGG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    1897 CCAGTTCATATCTCATGAGGGGTACATGATGAGGAAGTAAGCTGATTGGAAGTGCCTCAGGGGTTGG 1956

y     301 TAATATTCTACAATTGGGAAGCAGATAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    1957 TAATATTCTACAATTGGGAAGCAGATAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

y     361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

y     421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2136

y     481 GAAAGAGTTGAATGACTGGCTAACAACAAGAGAACAAGAACAAGAAAATGGAGAAGA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2137 GAAAGAGTTGAATGACTGGCTAACAACAAGAGAACAAGAACAAGAAAATGGAGAAGA 2196

y     541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256

y     601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCAATTCTCACATGGTGGTAGT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCAATTCTCACATGGTGGTAGT 2316

y     661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAGAACAACTTAAGGTATTGGG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAAGAACAACTTAAGGTATTGGG 2376

y     721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 2436

y     781 CCTTCCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2437 CCTTCCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 2496

y     841 AAAAGAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2497 AAAAGAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 2556

y     901 ATCACAGTCTTCAAAATGCGCCGTTTAAAGCGGATCTAGAAAGAAGAATCCAT  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2557 ATCACAGTCTTCAAAATGCGCCGTTTAAAGCGGATCTAGAAAGAAGAATCCAT 2616

y     961 GGGCAAACTGTATTCTCAAACAAGATCTCTTCTTCACACTGAAGAATAAGTCAGTGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2617 GGGCAAACTGTATTCTCAAACAAGATCTCTTCTTCACACTGAAGAATAAGTCAGTGAC 2676

y    1021 CCAGAAGACCGGAACCGGAACATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2677 CCAGAAGACCGGAACCGGAACATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2736

y    1081 ACTTGAAAAGAGTACAGCAGACTCATAGATTACTGGACAAGTTCCCCTGGACCTTGGA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2737 ACTTGAAAAGAGTACAGCAGACTCATAGATTACTGGACAAGTTCCCCTGGACCTTGGA 2796

y    1141 AAAGTTCTTGCCTGCTTACAGAAGCTGA 1170
          ||||||   ||||||
b    2797 TGAGCTGGACCTCAGCTGCGCGCCAAGCTGA 2826
```

Qy   181 TGCTGAGGACACATTGCAAGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1279 TGCTGAGGACACATTGCAAGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATACTCATGAGGGTACATGATGGTACATGATGGTTGCAGCCCATCAGGGCCGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1339 CCAGTTTCATACTCATGAGGGTACATGATGGTACATGATGGTTGCAGCCCATCAGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACACGGAAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACACGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCCTCAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1459 AACTGAAGTACAAGAGCAGATGAATCCTCAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAACAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1519 TAGCATGGAAAAACAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1578

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAAATGAGGAGGAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAAATGAGGAGGAGA 1638

Qy   541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1698

Qy   601 AGAGATCTAGAACAAGAACAAGTCAGGTCAATTCTTCACTCACATGGTGGTGGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1699 AGAGATCTAGAACAAGAACAAGTCAGGTCAATTCTTCACTCACATGGTGGTGGTAGT 1758

Qy   661 TGATGAATCTAGTGGAGATCGAGATCACGGCAACTGCTTGGAAGAACAACTTAAGTATTGGG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1759 TGATGAATCTAGTGGAGATCGAGATCACGGCAACTGCTTGGAAGAACAACTTAAGTATTGGG 1818

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1998

Qy   901 ATCAAGTCTTCAAAAACTGGCTTAAAGATCTAGAAAAGAAAAGAGAAAAAGCAATCCAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1999 ATCAAGTCTTCAAAAACTGGCTTAAAGATCTAGAAAAGAAAAGAGAAAAAGCAATCCAT 2058

Qy   961 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2059 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGTTGGGATAATTAGTCCAAAA 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGTTGGGATAATTAGTCCAAAA 2178

Qy 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
          |||||||||||||||||||||||||
Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 14
US-10-149-736-47
Sequence 47, Application US/10149736

Qy 1759 TGATGAATCTAGTGGAGATCGAGATCACGGCAACTGCTTGGAAGAACAACTTAAGTATTGGG 1818
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGTT 1998

Qy   901 ATCAAGTCTTCAAAAACTGGCTTAAAGATCTAGAAAAGAAAAGAGAAAAAGCAATCCAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1999 ATCAAGTCTTCAAAAACTGGCTTAAAGATCTAGAAAAGAAAAGAGAAAAAGCAATCCAT 2058

Qy   961 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2059 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGTTGGGATAATTAGTCCAAAA 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGTTGGGATAATTAGTCCAAAA 2178

Qy 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
          |||||||||||||||||||||||||
Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

ESULT 13
S-10-149-736-44
Sequence 44, Application US/10149736
S-10-149-736-44

Query Match   52.5%; Score 1103.4; DB 15; Length 11443;
Best Local Similarity 99.9%; Pred. No. 6.6e-310;
Matches 1104; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SEQ ID NO 47
LENGTH: 12057
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-149-736-47

Query Match   52.5%; Score 1103.4; DB 15; Length 12057;
Best Local Similarity 99.9%; Pred. No. 6.8e-310; Indels 1; Gaps 0;
Matches 1104; Conservative 0; Mismatches 0;

```
Qy     1 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGAGCCCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGAGCCCATT 1158

Qy    61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 TGAAGTAAACCTGGACCGTTATCATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 TGCTGAGGACACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTTCATCTCATGAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTTCATCTCATGAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATATCAGAAGATGAAGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGCAGATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459 AACTGAAGTACAAGAGCAGCAGATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAAACAAAGCAATTACATAGAGTTTTAATGCATCCCAGAATCAGAAACT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1519 TAGCATGGAAAAAACAAAGCAATTACATAGAGTTTTAATGCATCCCAGAATCAGAAACT 1578

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAACAAGGAAAAATGGGAGGAAGA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAACAAGGAAAAATGGGAGGAAGA 1638

Qy   541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1698

Qy   601 AGAAGATCTAGAACAAGAAGCAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTAGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699 AGAAGATCTAGAACAAGAAGCAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTAGT 1758

Qy   661 TGATGAATCTAGTGGAGAATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1759 TGATGAATCTAGTGGAGAATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACGAGACCGCTGGGTTCTTTTACAAGACAT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819 AGATCGATGGGCAAACATCTGTAGATGGACGAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1998

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058

Qy   961 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2059 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy  1021 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGGATAATTAGTCCAAAA 2178

Qy  1081 ACTTGAAAAGAGTACAGCACACAGACT 1105
           ||||||||||||||||||||||||||
Db  2179 ACTTGAAAAGAGTACAGCACACAGATT 2203
```

RESULT 15
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US2002010273lA1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-22

Query Match   52.5%; Score 1103.4; DB 9; Length 13957;
Best Local Similarity 99.9%; Pred. No. 7.5e-310; Indels 1; Gaps 0;
Matches 1104; Conservative 0; Mismatches 0;

```
Qy     1 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy    61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 TGAAGTAAACCTGGACCGTTATCATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTTCATCTCATGAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTTCATCTCATGAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATATCAGAAGATGAAGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTTCATCTCATGAGGGGTACTCAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 1398
```

```
b   1399  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  1458

y    361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1459  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  1518

y    421  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1519  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  1578

y    481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAACAAGGAAAATGGAGGAAGA  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1579  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAACAAGGAAAATGGAGGAAGA  1638

y    541  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAGTACAACAACATAAGGTGCTTCA  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1639  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAGTACAACAACATAAGGTGCTTCA  1698

y    601  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCCACATGGTGGTGTAGT  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1699  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCCACATGGTGGTGTAGT  1758

y    661  TGATGAATCTAGTGGAGATCACGCAACTCGCTGCTTTGGAAGAACAACTTAAGGTATTGGG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1759  TGATGAATCTAGTGGAGATCACGCAACTCGCTGCTTTGGAAGAACAACTTAAGGTATTGGG  1818

y    721  AGATCGATGGGCAAACATCTGTAGTGACAGAGACCGCTGGGTTCTTTTACAAGACAT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1819  AGATCGATGGGCAAACATCTGTAGTGACAGAGACCGCTGGGTTCTTTTACAAGACAT  1878

y    781  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGCTTTCAGA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1879  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  1938

y    841  AAAGAAGATGCAGTGAACAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1939  AAAGAAGATGCAGTGAACAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT  1998

y    901  ATCAAGTCTTCAAAAACTGCCGTTTAAAGCCGATCAGAAAAAGAAAAAGCAATCCAT  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1999  ATCAAGTCTTCAAAAACTGCCGTTTAAAGCCGATCAGAAAAAGAAAAAGCAATCCAT  2058

y    961  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2059  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC  2118

y   1021  CCAGAAGACCGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGGATAATTTAGTCCAAAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2119  CCAGAAGACCGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGGATAATTTAGTCCAAAA  2178

y   1081  ACTTGAAAAGAGTACAGCACAGACT  1105
          |||||||||||||||||||||||||
b   2179  ACTTGAAAAGAGTACAGCACAGATT  2203
```

earch completed: April 5, 2004, 16:32:49
ob time : 541.455 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

ум nucleic - nucleic search, using sw model

un on:    April 4, 2004, 11:55:32 ; Search time 542.141 Seconds
                                     (without alignments)
                               16463.377 Million cell updates/sec

itle:          US-09-845-416-2_COPY_900_3000
erfect score:  2101
equence: 1 gagctatgcctacacacagg........acctcagcactctggaagac 2101

coring table:  IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

earched:       3373863 seqs, 2124099041 residues

otal number of hits satisfying chosen parameters:   6747726

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :  N_Geneseq.29Jan04:*
        1:  geneseqn1980s:*
        2:  geneseqn1990s:*
        3:  geneseqn2000s:*
        4:  geneseqn2001as:*
        5:  geneseqn2001bs:*
        6:  geneseqn2002s:*
        7:  geneseqn2003as:*
        8:  geneseqn2003bs:*
        9:  geneseqn2003cs:*
       10:  geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	Query Match %	Length	DB	ID	Description
1	2101	100.0	4182	6	AAD37230	Aad37230 Human dys
2	2101	100.0	5149	6	AAD37255	Aad37235 Adeno-ass
3	1579.8	75.2	3999	6	AAD37234	Aad37234 Human dys
4	1579.8	75.2	4966	6	AAD37256	Aad37256 Adeno-ass
5	1579.8	75.2	4990	6	AAD37262	Aad37262 Adeno-ass
6	1443	68.7	3858	6	AAD37237	Aad37237 Human dys
7	1443	68.7	4825	6	AAD37257	Aad37257 Adeno-ass
8	1443	68.7	4848	6	AAD37263	Aad37263 Adeno-ass
9	1443	68.7	5060	6	AAD37264	Aad37264 Adeno-ass
10	1293	61.5	8689	6	ABK82000	Abk82000 DNA encod
11	1131	53.8	5952	5	AAD06794	Aad06794 Human dys
12	1114	53.0	4414	6	AAD37260	Aad37260 Adeno-ass
13	1103.4	52.5	5417	6	ABK81997	Abk81997 DNA encod
14	1103.4	52.5	11241	6	ABK82005	Abk82005 cDNA enco
15	1103.4	52.5	11443	6	ABK82002	Abk82002 DNA encod
16	1103.4	52.5	12923	1	AAN90338	Aan90338 Sequence
17	1103.4	52.5	13957	6	ABK81959	Abk81959 cDNA enco
18	1103.4	52.5	13957	6	ABT10904	Abt10904 Human bre
19	1103.4	52.5	13957	6	ABN95786	Abn95786 Gene #228
20	1103.4	52.5	13957	6	ABS69900	Abs69900 Human dys
21	1103.4	52.5	13957	6	ABS70403	Abs70403 Human bon
22	1102	52.5	3446	6	AAD37242	Aad37242 Human dys
23	1100.2	52.4	11058	6	AAD37229	Aad37229 Human dys
24	1097.8	52.3	1991	6	AAD37231	Aad37231 Human dys
25	999	47.5	2169	6	AAD37232	Aad37232 Human dys
26	999	47.5	3531	6	AAD37238	Aad37238 Human dys
27	999	47.5	4498	6	AAD37258	Aad37258 Adeno-ass
28	997	47.5	5339	6	ABK81998	Abk81998 DNA encod
29	996	47.4	5462	6	ABK81999	Abk81999 DNA encod
30	911.6	43.4	4402	3	AAZ48568	Aaz48568 A rod sho
31	866.6	41.2	13815	6	ABK81960	Abk81960 cDNA enco
32	866.6	41.2	13815	6	ABI99799	Abi99799 Mouse isc
33	866.6	41.2	19307	2	AAT27558	Aat27558 Shuttle v c
34	858.6	40.9	13815	6	AAV18885	Aav18885 Mus muscu
35	855.4	40.7	3275	1	AAN97129	Aan97129 Partial s
36	787	37.5	3510	6	AAD37240	Aad37240 Human dys
37	787	37.5	4476	6	AAD37259	Aad37259 Adeno-ass
38	777	37.0	1667	6	AAD37235	Aad37235 Human dys
39	718.8	34.2	4075	3	AAZ48569	Aaz48569 A rod sho
40	714.6	34.0	4402	3	AAZ48567	Aaz48567 A rod sho
41	652	31.0	1821	6	AAD37241	Aad37241 Human dys
42	450	21.4	1340	6	AAD37239	Aad37239 Human dys
43	409.2	19.5	3747	3	AAZ48566	Aaz48566 A rod sho
44	387	18.4	387	6	ABK81990	Abk81990 Human dys
45	361.4	17.2	9195	7	ACD19399	Acd19399 cDNA enco

ALIGNMENTS

```
RESULT 1
AAD37230
ID   AAD37230 standard; DNA; 4182 BP.
XX
AC   AAD37230;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Human dystrophin minigene delta4173.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 43-44; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
```

```
:C   dystrophy (BMD) in a mammalian subject. The present sequence is human
:C   dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
:C   hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
:X   hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
:Q   Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;

Query Match           100.0%;  Score 2101;   DB 6;   Length 4182;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 2101;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
```

```
Qy      1  GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTGACCTGACCACGGAGCCCATT    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900  GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTGACCTGACCACGGAGCCCATT   959

Qy     61  TCCTTCACAGCAGTTGGAAGCTCCTAGAGACAAGTCATTGGCAGTTCATTGATGGAGAG   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    960  TCCTTCACAGCAGTTGGAAGCTCCTAGAGACAAGTCATTGGCAGTTCATTGATGGAGAG  1019

Qy    121  TGAAGTAAACCTGGACCCGTTATCAAGCACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020  TGAAGTAAACCTGGACCCGTTATCAAGCACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC  1079

Qy    181  TGCTGAGGACACATTGCAAGCACAAGGAGATTCTAATGATGTGGAAGTGGTGAAAGA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080  TGCTGAGGACACATTGCAAGCACAAGGAGATTCTAATGATGTGGAAGTGGTGAAAGA  1139

Qy    241  CCAGTTCATCTGAGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140  CCAGTTCATCTGAGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG  1199

Qy    301  TAATATTCTACAATTGGGAAGTAAGCTGATTGGGAACAGGAAAATTATCAGAGATGAAGA   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200  TAATATTCTACAATTGGGAAGTAAGCTGATTGGGAACAGGAAAATTATCAGAGATGAAGA  1259

Qy    361  AACTGAAGAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGCTGAATGCTGCCAGGGTAGC   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260  AACTGAAGAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGCTGAATGCTGCCAGGGTAGC  1319

Qy    421  TAGCATGGAAAAACAAGCAATTTACATGAGTTTAATGCATCCCAGAATCAGAAACT   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320  TAGCATGGAAAAACAAGCAATTTACATGAGTTTAATGCATCCCAGAATCAGAAACT  1379

Qy    481  GAAAGAGTTGATGATGAACTGGCTAACAAAAACAGAAGAACAAGGAAAATGGAGGAGA   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380  GAAAGAGTTGATGATGAACTGGCTAACAAAAACAGAAGAACAAGGAAAATGGAGGAGA  1439

Qy    541  GCCTCTGGACCTGATCTTGAAGACCTTAAAACGCCAAGTCAATTCTCACTCATAAGGTGCTTCA   600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440  GCCTCTGGACCTGATCTTGAAGACCTTAAAACGCCAAGTCAATTCTCACTCATAAGGTGCTTCA  1499

Qy    601  AGGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCCATCATAAGGTGATTGGG   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500  AGGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCCATCATAAGGTGATTGGG  1559

Qy    661  TGATGAATCTAGTGGAAGATCGTGAGTGATGAACTTAAGGTATTGGG   720
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1560  TGATGAATCTAGTGGAAGATCGTGAGTGATGAACTTAAGGTATTGGG  1619

Qy    721  AGATCGATGGGCAAACATCTGTAGATGGAACAGAGACCGCTGGGTTCTTTTACAAGACAT   780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1620  AGATCGATGGGCAAACATCTGTAGATGGAACAGAGACCGCTGGGTTCTTTTACAAGACAT  1679

Qy    781  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA   840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1680  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  1739

Qy    841  AAAAGAAGATGCAGTGAACGTCACACAACTGCTTAAAGATCAAAAATGAAATGTT   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1740  AAAAGAAGATGCAGTGAACGTCACACAACTGCTTAAAGATCAAAAATGAAATGTT  1799

Qy    901  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1800  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT  1859
```

```
Db   1800  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT  1859

Qy    961  GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1860  GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  1919

Qy   1021  CCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1920  CCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  1979

Qy   1081  ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGACCTGA  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1980  ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGACCTGA  2039

Qy   1141  AAAGTTCTTGCCTGGCTGAAACAACTGCCAATGCCTACAGGATGCTAC  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2040  AAAGTTCTTGCCTGGCTGAAACAACTGCCAATGCCTACAGGATGCTAC  2099

Qy   1201  CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2100  CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA  2159

Qy   1261  AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2160  AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAAACAG  2219

Qy   1321  CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCAGTCCGTGTTACAAAGACGTTT  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2220  CCAAAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCAGTCCGTGTTACAAAGACGTTT  2279

Qy   1381  GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGTCCCA  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2280  GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGTCCCA  2339

Qy   1441  TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2340  TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT  2399

Qy   1501  GTGGCTACAGCTGAAAGATGATGATGAATTAAGCCGGAGGCACCTATTGGAGGGGACTTTCC  1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2400  GTGGCTACAGCTGAAAGATGATGATGAATTAAGCCGGAGGCACCTATTGGAGGGGACTTTCC  2459

Qy   1561  AGCAGTTCAGAAGCAGAAGCAGAGCATGTACATAGGGCTTCAAGGAATTGAAAACTAAAGA  1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460  AGCAGTTCAGAAGCAGAAGCAGAGCATGTACATAGGGCTTCAAGGAATTGAAAACTAAAGA  2519

Qy   1621  ACCTGTAATCATGAGTACTCTTGAGACTGTACGGACTGTACGAATATTCTGACAGAGCTG  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2520  ACCTGTAATCATGAGTACTCTTGAGACTGTACGGACTGTACGAATATTCTGACAGAGCTG  2579

Qy   1681  AGGACTAGAAGCAGAACTCTACCAGGAGCCCAAGCTGAGGAGGCCAGAGAGCCCAGAA  1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2580  AGGACTAGAAGCAGAACTCTACCAGGAGCCCAAGCTGAGGAGGCCAGAGAGCCCAGAA  2639

Qy   1741  TGTCACTCGGCTTCTACGGAAAGCAGGCTCAATACTGAGTGGGAAAATTGAA  1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2640  TGTCACTCGGCTTCTACGGAAAGCAGGCTCAATACTGAGTGGGAAAATTGAA  2699

Qy   1801  CCTGCACTCCGCTGGCTGGCACGGATGAGCTGAAAAGATGAGACCCTTGAAAGACTCCAGGAACT  1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700  CCTGCACTCCGCTGGCTGGCACGGATGAGCTGAAAAGATGAGACCCTTGAAAGACTCCAGGAACT  2759

Qy   1861  TCAAGAGGCCACGGGATGAGCTGGCGCCAAGCTGGACCTGAGGTGATCAAGGGATC  1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2760  TCAAGAGGCCACGGGATGAGCTGGCGCCAAGCTGGACCTGAGGTGATCAAGGGATC  2819

Qy   1921  CTGGCAGCCCGTGGCGCCGATCTCCTCATTGACTCTCTCCAAGATCAGCTCGAGAAAGTCAA  1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2820  CTGGCAGCCCGTGGCGCCGATCTCCTCATTGACTCTCTCCAAGATCAGCTCGAGAAAGTCAA  2879

Qy   1981  GGCACTTCGAGGAGAATTGCGCCTTCTGAAAGAAGCGTGAGCCACGTCAATGACCTTGC  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2880  GGCACTTCGAGGAGAATTGCGCCTTCTGAAAGAAGCGTGAGCCACGTCAATGACCTTGC  2939
```

```
y   2041 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCACTCTGGAAGA 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2940 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCACTCTGGAAGA 2999

y   2101 C 2101
         |
b   3000 C 3000
```

RESULT 2
AAD37255

```
D  AAD37255 standard; DNA; 5149 BP.
X  AAD37255;
C  21-AUG-2002  (first entry)
B  Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
X  Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
X  adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
X  Becker muscular dystrophy; ds.
S  Homo sapiens.
S  Unidentified.
S  Chimeric.
N  WO200183695-A2.
O  08-NOV-2001.
X  27-APR-2001; 2001WO-US013677.
X  28-APR-2000; 2000US-0200777P.
A  (XIAO/) XIAO X.
I  Xiao X;
R  WPI; 2002-049342/06.
F  New dystrophin minigene for treating Duchenne or Becker muscular
   dystrophy. The minigene comprises an N-terminal domain or modified N-
   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
   domains and cysteine-rich domain of dystrophin or utrophin genes. The
   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
   gene.
X  Example 1; Page 57-59; 71pp; English.
C  The present invention relates to an isolated nucleotide sequence encoding
   a dystrophin minigene. The minigene comprises N-terminal or modified N-
   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
   domains and cysteine-rich domains of dystrophin or utrophin genes. The
   invention also relates to a recombinant adeno-associated virus (AAV). The
   comprising dystrophin minigene operably linked to an expression control
   element. The dystrophin minigene in operable linkage with an expression
   control element, in a recombinant adeno-associated virus or retrovirus is
   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
   vector plasmid construct containing human dystrophin minigene, a muscle
   creatine kinase (MCK) promoter and a small polyA signal sequence.
Q  Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;
```

```
Query Match             100.0%;  Score 2101;  DB 6;  Length 5149;
Best Local Similarity   100.0%;  Pred. No.0;
Matches 2101   Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Y      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT 60
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
b   1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTTCTCACCCTACACGGAGCCCATT 1716

Y     61 TCCTTCACAGCATTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAGAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1717 TCCTTCACAGCATTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAGAG 1776

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC 1836

Qy   181 TGCTGAGGACACATTGCAAGCACAGCAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1837 TGCTGAGGACACATTGCAAGCACAGCAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy   241 CCAGTTTCATACTCATGAGGGGTACATGATGGGGTACAGCCCATCAGGGCCGGGTTGG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1897 CCAGTTTCATACTCATGAGGGGTACATGATGGGGTACAGCCCATCAGGGCCGGGTTGG 1956

Qy   301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1957 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC 2076

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 2136

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAATGAGAGGAAGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAATGAGAGGAAGA 2196

Qy   541 GCCTCCTTGACCTGATCTTGAAGACTACAAGCCAAGTACACAACAAGCATAAGGTGCTTCA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2197 GCCTCCTTGACCTGATCTTGAAGACTACAAGCCAAGTACACAACAAGCATAAGGTGCTTCA 2256

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCCACATGGTGTGGTAGT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCCACATGGTGTGGTAGT 2316

Qy   661 TGATGAATCTAGTGGAGATCACCGCAACTGCTGTTTGGAGAAGAACAACTAAGGTATTGGG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2317 TGATGAATCTAGTGGAGATCACCGCAACTGCTGTTTGGAGAAGAACAACTAAGGTATTGGG 2376

Qy   721 AGATCGATGGGCAAACATCGTAGATGGACAGAAGACCGCTGGGTTCTTTAAAGATCAAAATGAAATGTT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2377 AGATCGATGGGCAAACATCGTAGATGGACAGAAGACCGCTGGGTTCTTTAAAGATCAAAATGAAATGTT 2436

Qy   781 CCTTCCAAATGGCCAAAATGGCAACGTCTTACTGTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2437 CCTTCCAAATGGCCAAAATGGCAACGTCTTACTGTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 2496

Qy   841 AAAAGAAGATGAGCAGTGAACAAGATTCACACAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2497 AAAAGAAGATGAGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 2556

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAGAAAAAGAAATCCAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2557 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAGAAAAAGAAATCCAT 2616

Qy   961 GGGCAAACTGTATTCCTCAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2617 GGGCAAACTGTATTCCTCAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 2676

Qy  1021 CCAGAAGACGGAAGCATGGCGATGAATAACTTTGCCCGGTGTGTTGGGATAATTAGTCCAAAA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2677 CCAGAAGACGGAAGCATGGCGATGAATAACTTTGCCCGGTGTGTTGGGATAATTAGTCCAAAA 2736

Qy  1081 ACTTGAAAGAGTACAGACAGACTCATAGATTACTGCAACAGTTCCCCTGACCTGGA 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2737 ACTTGAAAGAGTACAGACAGACTCATAGATTACTGCAACAGTTCCCCTGACCTGGA 2796

Qy  1141 AAAGTTTCTTGCCTGGCTTACAGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAGAG 1200
```

```
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 46-47; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene; H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;

Query Match             75.2%;  Score 1579.8;  DB 6;  Length 3999;
Best Local Similarity   87.9%;  Pred. No. 0;
Matches 1846;  Conservative   0;  Mismatches  72;  Indels 183;  Gaps   4;

Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACCGGAGCCGCCATT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACCGGAGCCGCCATT 959

Qy    61 TCCTTCACAGCATTTGGAAGTCCTGGAAGACAAGTCATTGGCAGTCATTGGTGCTTCTTTC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTTGGAAGTCCTGGAAGACAAGTCATTGGCAGTCATTGGTGCTTCTTTC 1019

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGGAGAAGAAGTTTAGAACAGCTTCTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGGAGAAGAAGTTTAGAACAGCTTCTTTC 1079

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1080 TGCTGAGGACACATTGCAAGCACAAGGAGATTTCAATGATGTGGAAGTGGTGAAAGA 1139

Qy   241 CCAGTTCATACTCATCATGATGAGGGTACATGATGATGCCCCATCAGGGCCGGGTTGG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1140 CCAGTTTCATACTCATCATGATGAGGGTACATGATGATGCCCCATCAGGGCCGGGTTGG 1199

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAATTATCATCGAGAGATGAAGA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  2797 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 2856
Qy  1201 CCGTAAGAAGAGGCTCCTAGAGACTTCCAAGGAGTAAAAGAGAGCTGATGAATGAAACAATGGCA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2857 CCGTAAGAAGAGGCTCCTAGAGACTTCCAAGGAGTAAAAGAGAGCTGATGAACAATGGCA 2916
Qy  1261 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAG 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2917 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAG 2976
Qy  1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2977 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT 3036
Qy  1381 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3037 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3096
Qy  1441 TTTGGAGCCAGTTCTGACCAGTGGAAGCCGTCTGACACCTTTCTCTGCAGGAACTTCTGT 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3097 TTTGGAGCCAGTTCTGACCAGTGGAAGCCGTCTGACACCTTTCTCTGCAGGAACTTCTGT 3156
Qy  1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCCACCTATTGGAGGCGACTTTCC 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3157 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCCACCTATTGGAGGCGACTTTTC 3216
Qy  1561 AGCAGTTCAGAAGCAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAAAACTAAAGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3217 AGCAGTTCAGAAGCAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAAAACTAAAGA 3276
Qy  1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTTCTGACAGAGCAGCCTTTGGA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3277 ACCTGTAATCATGAGTACTCTTGAGACTACTCTTGAGACTGTACGACAGAGCAGCCTTTGGA 3336
Qy  1681 AGGACTAGAAACTCTACCAGGAGCCCAGAGAGCTGCTCCTCGAGAGAGAGCAGCCCAGAA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3337 AGGACTAGAAACTCTACCAGGAGCCCAGAGAGCTGCTCCTCGAGAGAGAGCAGCCCAGAA 3396
Qy  1741 TGTCACTCGGCTTCACGAAAGCAGCAGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3397 TGTCACTCGGCTTCACGAAAGCAGCAGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3456
Qy  1801 CCTGCACTCCGCTGACTGGCAGGAGAAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3457 CCTGCACTCCGCTGACTGGCAGGAGAAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 3516
Qy  1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3517 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 3576
Qy  1921 CTGCAGCCGTGGGCGATCTCCTCATTGACCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3577 CTGCAGCCGTGGGCGATCTCCTCATTGACCTCTCCAAGATCACCTCGAGAAAGTCAA 3636
Qy  1981 GGCACTTCGAGAGGAGAAATTGCGCGCCTCTGAAAGACGTGAGCCACGTCAATGACCTTGC 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3637 GGCACTTCGAGAGGAGAAATTGCGCGCCTCTGAAAGACGTGAGCCACGTCAATGACCTTGC 3696
Qy  2041 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCAGCTCTCACCGTATAACCTCCAGCAGCTCTGGAAGA 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3697 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCAGCTCTCACCGTATAACCTCCAGCACTCTGGAAGA 3756
Qy  2101 C 2101
         |
Db  3757 C 3757
```

RESULT 3
AAD37234
```
ID   AAD37234 standard; DNA; 3999 BP.
XX
AC   AAD37234;
XX
```

```
1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACACAGAAAATTATCGAGAGATGAAGA 1259

 361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1319

 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1379

 481 GAAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAATGAGGAAGA 540
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1380 GAAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAATGAGGAAGA 1439

 541 GCCTCTTGGACCTGATCTTGAAGACTACCAAGTACAACATAAGGTGCTTCA 600
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1440 GCCTCTTGGACCTGATCTTGAAGACTACCAAGTACAACATAAGGTGCTTCA 1499

 601 AGAAGATCTAGAAACAGAACAAGTCAGGGTCAATTCTCCTCACTCGTGGTGGTAGT 660
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
1500 AGAAGATCTAGAAACAGAACAAGTCAGGGTCAATTCTCCTCACTCGTGGTGGTAGT 1559

 661 TGATGAATCTAGTGGAGAATCAGTCAACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 720
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1560 TGATGAATCTAGTGGAGAATCAGTCAACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1619

 721 AGATCCATGGCAAACATCTGTAGATGACAGAAGACCGTCGGGTTCTTTACAAGACAT 780
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1620 AGATCCATGGCAAACATCTGTAGATGACAGAAGACCGTCGGGTTCTTTACAAGACCA 1679

 781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCCTTTTAGTGCATGGCTTTCAGA 840
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
1680 CTAGCTCCTGGACTGACTACCAACTGGTTAAAGATCAAAATGAAATGTT 1727

 841 AAAAGAAGATGCAGTGAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 900
     |||||||||||||||||||||||||||||||
1728 GACTGTTACTCTGGTGACACA------ACCTGTGGTTACTAAGGAAACTGCCATCT- 1777

 901 ATCAAGTCTTCAAAAACTGCGGCCGTTTTAAAAGCGGATCTAGAAAGAGAAAAGCAATCCAT 960
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1778 --------------------------------------------------------------- 1777

1778 CCAAACTAGAAAATGCCATCTTCCTTGGGGATAATTAGTCCAAAA 1812

 961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
     ||||||||||||||||||||||||||
1778 CCAAACTAGAAAATGCCATCTTCCTTGGGGATAATTAGTCCAAAA 1812

1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGGTGTTGGGATAATTAGTCCAAAA 1080
     |||||||||||||||||||||
1813 ----------------------------------------------------------- 1812

1081 ACTTGAAAAGAGTACAGACAGACTCATAGATTACTGCAACAGTTCCCCTGACCTGGA 1140
     |||||||||||||||||
1813 ------GTACCTACTACTGTAGAGATTACTGCAACAGTTCCCCTGACCTGGA 1856

1141 AAAGTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1200
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1857 AAAGTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1916

1201 CCGTAAGGAAAGGCCTCCAGGAGAGCTAGAGACTGCCAAGGGAGTAAAGAGCTGATGAAACAATGGCA 1260
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1917 CCGTAAGGAAAGGCCTCCAGGGAGACTCCAGGAGAGCTGATGAATGAAACAACAATGGCA 1976

1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1977 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2036

1321 CCAAAAAATCCTGAATCCTGAGAATCCTGAGTCGAGTGAACTTGTTACAAAGACGTTT 1380
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2037 CCAAAAAATCCTGAATCCTGAGAATCCTGAGTCGAGTGAACTTGTTACAAAGACGTTT 2096

1381 GGATAACATGAACTTCAAGTGGAGTGAACTTCTCTCAACATTAGGTCCCA 1440
     |||||||||||||||||||||||||||||||||||||||||||||||||
2097 GGATAACATGAACTTCAAGTGGAGTGAACTTCTCTCAACATTAGGTCCCA 2156
```

```
1441 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTCGCACCTTTCTCTGCGAGGAACTTCTGT 1500
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2157 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTCGCACCTTTCTCTGCGAGGAACTTCTGGT 2216

1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCCGGCAGGCACCTATTGGAGGGCACTTTCC 1560
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2217 GTGGCTACAGCTGAAAGATGATGAATTAAGCCCGGCAGGCACCTATTGGAGGGCACTTTCC 2276

1561 AGCAGTTCAGAGACAGACGATGATGAACAGAGATGTACATAGGGGCCTTCAGAGGGAATTGAAAACTAAAGA 1620
     |||||||||||||||||||||||||||||||||||||||||||||||||||||
2277 AGCAGTTCAGAGACAGACGATGATGAACAGAGATGTACATAGGGGCCTTCAGAGGGAATTGAAAACTAAAGA 2336

1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACAGAATATTTCTGACAGAGCCTTTGGA 1680
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2337 ACCTGTAATCATGAGTACTCTTGAGACTGTACAGAATATTTCTGACAGAGCCTTTGGA 2396

1681 AGGACTAGAGAAAACTTCACCAGGAGCCCAGAGAGCTGCTCCTCGAGGAGAGCCCAGAA 1740
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2397 AGGACTAGAGAAAACTTCACCAGGAGCCCAGAGAGCTGCTCCTCGAGGAGAGCCCAGAA 2456

1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTCGGGAAAAATTGAA 1800
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2457 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTCGGGAAAAATTGAA 2516

1801 CCTGCACTCCGCTGACTGGCTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGACT 1860
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2517 CCTGCACTCCGCTGACTGGCTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 2576

1861 TCAAGAGGGCCACGGATGAGCTGGACCTCAAGCTGAGGTGATCAAGGGATC 1920
     |||||||||||||||||||||||||||||||||||||||||||||||||||
2577 TCAAGAGGGCCACGGATGAGCTGGACCTCAAGCTGAGGTGATCAAGGGATC 2636

1921 CTGGCAGCCGCGTGGGCGATCTCCTCATTGACTCTCTCGAGATCACCTCGAGAAAGTCAA 1980
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2637 CTGGCAGCCGCCCGTGGGCGATCTCCTCATTCAGCTCTCTCGAGATCACCTCGAGAAAGTCAA 2696

1981 GGCACTTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCGAGCCACGTCAATGACCTTGC 2040
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2697 GGCACTTCGAGGAGAAATTGCGCCTTGGGCATTCAGCTTGCGCCTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2756

2041 TCGCCAGCTTACCACTTTGGGCATTCAGCTTGCGTATAACCTCAGCACTCTGGAAGA 2100
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2757 TCGCCAGCTTACCACTTTGGGCATTCAGCTTGCGTATAACCTCAGCTCTGGAAGA 2816

2101 C 2101

2817 C 2817
```

CX 28-APR-2000; 2000US-0200777P.
PR (XIAO/) XIAO X.
PA (XIAO/) XIAO X.
PI Xiao X;
XA Xiao X;
DR WPI; 2002-049342/06.
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
PS Example 1; Page 59-60; 71pp; English.
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;

Query Match   75.2%; Score 1579.8; DB 6; Length 4966;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative   0; Mismatches   72; Indels 183; Gaps 4;

Qy   1 GAGCTATGCTACCACACAGGCTGCTTATGTCACCACCCTCTGACCCTACACGGAGCCATT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1657 GAGCTATGCTACCACACAGGCTGCTTATGTCACCACCCTCTGACCCTACACGGAGCCATT 1716

Qy   61 TCCTTCACAGCATTTGGAAGCTTCTGAAGACAAGCAAGCAGTTTAGAAGAACAGCTTGG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1717 TCCTTCACAGCATTGGAAGCTCTCGAAGACAAGCAAGCAGTTTAGAAGAACAGCTTGG 1776

Qy 121 TGAAGTAAACCTGACCGTTATCAAACCAGCTTTAGAAGAAGAAGCAGCTTTAGAAGAAGAAGCA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1777 TGAAGTAAACCTGACCGTTATCAAACCAGCTTTAGAAGAAGAAGCAGCTTTAGAAGAAGAAGCA 1836

Qy 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy 241 CCAGTTCATACTCATGAGGGTACATGGGTACAGCCCATGACAGCCCATCAGGGCCGGGTTGG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1897 CCAGTTCATACTCATGAGGGTACATGGGTACAGCCCATGACAGCCCATCAGGGCCGGGTTGG 1956

Qy 301 TAATATTCTACAATTGGGAAGTAAGCTGATGAATGAACAGGAGAAAATTATCAGAAGATGAAGA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1957 TAATATTCTACAATTGGGAAGTAAGCTGATGAATGAACAGGAGAAAATTATCAGAAGATGAAGA 2016

Qy 361 AACTGAAGTACAAGAGAAAAAACAGAATCCAATTCACATAGAGTTTGGAATCTCCGATGCTCAGGGTAGC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2017 AACTGAAGTACAAGAGAAAAAACAGAATCCAATTCACATAGAGTTTGGAATCTCCGATGCTCAGGGTAGC 2076

Qy 421 TAGGCATGGAAATTGGAATGCTGCTAACAAGCAATTTACATAGGCGCAAAAAACAGAGAAGAAACAGAAACT 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2077 TAGGCATGGAAATTGGAATGCTGCTAACAAGCAATTTACATAGGCGCAAAAAACAGAGAAGAAACAGAAACT 2136

Qy 481 GAAAGAGTTGAATGACTGGCTACAAAAACAGAGAAGAACAAGAAGAAATGGAAGAAGA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2137 GAAAGAGTTGAATGACTGGCTACAAAAACAGAGAAGAACAAGAAGAAATGGAAGAAGA 2196

Qy 541 GCCTCTTGGACCTGATCTTGAAGACCTAGAAACGCCAAGTACAACACACATAAGGTGCTTCA 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTAGAAACGCCAAGTACAACACACATAAGGTGCTTCA 2256

Qy 601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGTGGTGGTAGT 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGTGGTGGTAGT 2316

Qy 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 2376

Qy 721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCTGGGTTCTTTTACAAGACAT 780
         |||||||||||||||||||||||||||||||||||||| |||||||||||||
Db 2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACCA 2436

Qy 781 CCTTCTCAAATGCAACGTCTTACTGAAGAACAGCAGTGCCTTTTAGTGCATGGCTTTCAGA 840
         ||||| ||||||||||||||||||
Db 2437 GCCTGAC------   CTAGCTCCTGACTGACCACTATTGGAGCCTCTCCTACTCA 2484

Qy 841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGCTTTAAAGATCAAAATGAAAATGTT 900
         |||| ||||
Db 2485 GACTGTTACTCGTGACACA------ACCTGTGGTTACTAAGGAAACTGCCATCT- 2534

Qy 901 ATCAAGTTCTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960

Db 2535 ------ 2534

Qy 961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACAACTGCTGAAGAATAAGTCAGTGAC 1020
                                 ||||| |||||
Db 2535 ------CCAAACTAGAAATGCCCATCTTCCTTGATGTTGGAG--- 2569

Qy 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 1080

Db 2570 ------ 2569

Qy 1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA 1140
                      |||||||||||||||||||||||||||||||||||
Db 2570 ------GTACTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA 2613

Qy 1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTCACAGGATGCTAC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2614 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTCACAGGATGCTAC 2673

Qy 1201 CCGTAAGGAAAGGCTCCTAGAAGACTTCCAAGGGAGTGAAAAAGAGCTGATGAAACAATGGCA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2674 CCGTAAGGAAAGGCTCCTAGAAGACTTCCAAGGGAGTGAAAAAGAGCTGATGAAACAATGGCA 2733

Qy 1261 AGACCTCCAAGGTGAAATTGAAGCTCCAACAGATGTTATCAACACCTGATGAAAAACAG 1320
         ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 2734 AGACCTCCAAGGTGAAATTGAAGCTCACAACAGATGTTATCAACACCTGATGAAAAACAG 2793

Qy 1321 CCAAAAAATCCTGAGTCCCTGGAAGGTGAACTGGAGTGGAGTCAGGGCCGGGTTGG 1380
         |||||||||||||||| |||||||||||| |||||||||||||
Db 2794 CCAAAAAATCCTGAGATCCCTGGAAGGTGAGTGGAGTCAGCCTGGTTACACAAGACGTTT 2853

Qy 1381 GGATAACATGAACTTCAAGTGACACCAGTGAGCGTCGCACCCTTCTCTCGCAGGAACTTCTGGT 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2854 GGATAACATGAACTTCAAGTGACACCAGTGAGCGTCGCACCCTTCTCTCGCAGGAACTTCTGGT 2913

Qy 1441 TTTGGAAGCCAGTTCGACCAGTGAAAGATGAATTAAGCCGGACCACCTATTGGGAGGCGACTTCC 1500

Db 2914 GTGGCTACAGCTGAAAGATGATGAAAGATGAATGAATGCCTCAGGGTAGC 2973

Qy 1501 GTGGCTACAGCTGAAAGATGATGAAAGATGAATGAATGCCTCAGGGTAGC 1560

Db 2974 AGCAGTTCAGAAGCAGCAGAACGATGTACATAGGCGCCTTCAAGAGGCACCTATTGGGAATTGAAAACTAAAGA 3033

Qy 1561 AGCAGTTCAGAAGCAGCAGAACGATGTACATAGGCGCCTTCAAGAGGCACCTATTGGGAATTGAAAACTAAAGA 1620

Db 3034 ACTTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 3093

Qy 1621 ACTTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 1680

Db 3094 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGA 3153

XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigene, a
CC   cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ   Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;

Query Match   75.2%; Score 1579.8; DB 6; Length 4990;
Best Local Similarity   87.9%; Pred. No. 0;
Matches 1846;   Conservative   0; Mismatches 72; Indels 183; Gaps 4;

```
Qy      1 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1740

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTCATTGATGAGGAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTCATTGATGAGGAG 1800

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGTGGCTTCTTTC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGTGGCTTCTTTC 1860

Qy    181 TGCTGAGGACACAATTGCAAGCAAGCACCAAGGAGATTTCTAATGATGTGAAGTGGTGAAAGA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCTGAGGACACAATTGCAAGCAAGCACCAAGGAGATTTCTAATGATGTGAAGTGGTGAAAGA 1920

Qy    241 CCAGTTCATACTCATGAGGGGTACATGAATTGACAGCCCATCAGGGGCCGGGTTGG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CCAGTTCATACTCATGAGGGGTACATGAATTGACAGCCCATCAGGGGCCGGGTTGG 1980

Qy    301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA 2040

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2100

Qy    421 TAGCATGGAAAAACAAAGCAATTACATTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TAGCATGGAAAAACAAAGCAATTACATTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2160

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGCCAAGACCAAGTACAACATAAGGTGCTTCA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GAAAGAGTTGAATGACTGGCTAACAAAAACAGCCAAGACCAAGTACAACATAAGGTGCTTCA 2220

Qy    541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACATAAGGTGCTTCA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACATAAGGTGCTTCA 2280

Qy    601 AGAAGATCTAGAACAAGAGAACAAGAAGAACAAGTCAGGGTCAATTCTCTCACTCGTGTGGTAGT  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 AGAAGATCTAGAACAAGAGAACAAGAAGAACAAGTCAGGGTCAATTCTCTCACTCGTGTGGTAGT 2340

Qy    661 TGATGAATCTAGTGGAATCTAGTGGGATCACCGCAACTGCTGCTTGGAAGAACAACTAAGGTATTGGG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TGATGAATCTAGTGGAATCTAGTGGGATCACCGCAACTGCTGCTTGGAAGAACAACTAAGGTATTGGG 2400

Qy    721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGTTCTTTACAAGACAT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGTTCTTTACAAGACCA 2460

Qy    781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTTAGTGCATGGCTTTCAGA  840
```

```
Qy   1681 AGGACTAGAGAAACTCTACCAGGAGCTGCCTCCTGAGGAGGAGCCCAGAA 1740
          ||||||||||||||||||||||||||||||||||||||
Db   3154 AGGACTAGAGAAACTCTACCAGGAGCCCAGGAGCCCAGAGAGCTGCCCCAGAA 3213

Qy   1741 TGTCACTCCGGCTTCACAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
          ||||||||||||||||||||||||||||||||||||||
Db   3214 TGTCACTCGGCTTCACAAGAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3273

Qy   1801 CCTGCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1860
          ||||||||||||||||||||||||||||||||||||||
Db   3274 CCTGCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 3333

Qy   1861 TCAAGAGGCCACGATGAGCTGACCTCAAGTTGCGCGAAGTGAGTGATCAAGGGATC 1920
          ||||||||||||||||||||||||||||||||||||||
Db   3334 TCAAGAGGCCACGATGAGCTGACCTCAAGTTGCGCGAAGTGAGTGATCAAGGGATC 3393

Qy   1921 CTGGCAGCGCCGTGGGCGGATCTCCATTGACTCTTCTCCAAGATCACCTCGAGAAAGTCAA 1980
          ||||||||||||||||||||||||||||||||||||||
Db   3394 CTGGCAGCCGCCGTGGGCGGATCTCCATTGACTCTTCTCCAAGATCACCTCGAGAAAGTCAA 3453

Qy   1981 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAAGAAGAACGTGAGCCCAGTCAATGACCTTGC 2040
          ||||||||||||||||||||||||||||||||||||||
Db   3454 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAAGAAGAACGTGAGCCACGTCAATGACCTTGC 3513

Qy   2041 TCGGCAGCTTACCACTTTGGGCATTCAGCTCTCACCCGTATAACCTCAGCACTTCTGGAAGA 2100
          ||||||||||||||||||||||||||||||||||||||
Db   3514 TCGGCAGCTTACCACTTTGGGCATTCAGCTCTCACCCGTATAACCTCAGCACTTCTGGAAGA 3573

Qy   2101 C 2101
          |
Db   3574 C 3574
```

RESULT 5
AD37262
CC   AAD37262 standard; DNA; 4990 BP.
XX
CC   AAD37262;
XX
FT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Cytomegalovirus.
OS   Unidentified.
OS   Chimeric.
XX
DR   WO200183695-A2.
XX
DR   08-NOV-2001.
XX
PN   27-APR-2001; 2001WO-US013677.
PF   28-APR-2000; 2000US-0200777P.
PR   (XIAO/) XIAO X.
AA   Xiao X;
XI   WPI; 2002-049342/06.
XR   New dystrophin minigene for treating Duchenne or Becker muscular
XT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
FT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
FT   gene.
XX
XS   Example 1; Page 67-68; 71pp; English.

```
Qy  2461 GCCTGAC----------CCAGCTCCTGGACTGACCACCATTGAGCCTCTCCTACTCA 2508
              ||  |         |||  ||   |||||||||||||||||||||||
Db  3358 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGCTGATCAAGGGATC 3417

Qy  841  AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAAGTGTT 900
             |||||||       ||||            |  |||   ||||   ||||
Db  1921 CTGGCAGCCCGTGGGGCGATCCTCCATTGACTCTCCCAAGATCACCTCCGAGAAAGTCAA 1980

Qy  2509 GACTGTTACTCTGGTGACACA--------ACCTGTGGTTACTAAGGAAACTGCCATCT- 2558
             ||    ||||   |||||              | ||| ||   ||||||||||||
Db  3418 CTGGCAGCCCGTGGGGCGATTCCTCCATTGACTCTCCCAAGATCACCTCCGAGAAAGTCAA 3477

Qy  901  ATCAAGTCTTCAAAAACTGGCCGTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 GGCACTTCGAGGAGAAATTGCCCTCTGAAAGAGAACGTGAGCCACGTCAAATGACCTTGC 2040

Qy  2559 ------------------------------------------- 2558
Db  3478 GGCACTTCGAGGAGAAATTGCGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3537

Qy  961  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
              ||| ||| | |||   | |    |  ||  | |||  || ||||  ||  ||
Db  2041 TCGCCAGCTTACCACCTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100

Qy  2559 ----------CCAAACTAGAAATGCCATCTTCCTGATGTTGGAG------- 2593
Db  3538 TCGCCAGCTTACCACCTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3597

Qy  1021 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db  2101 C 2101

Qy  2594 ------------------------------------------- 2593
Db  3598 C 3598
```

RESULT 6
AAD37237
ID   AAD37237 standard; DNA; 3858 BP.
XX
AC   AAD37237;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy. The minigene comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 48-49; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;

Query Match   68.7%;   Score 1443;   DB 6;   Length 3858;

```
Qy  1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2594 -----------GTACTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA 2637

Qy  1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGAGATGCTAC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2638 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGAGATGCTAC 2697

Qy  1201 CCGTAAGGAAAAGGCTCCTAGAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2698 CCGTAAGGAAAAGGCTCCTAGAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 2757

Qy  1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAG 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2758 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAG 2817

Qy  1321 CCAAAAAATCCTGAGATCCTGGAAGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT 1380
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2818 CCAAAAAATCCTGAGATCCCTGGAAGTTCCCATGATGCAGTCCTGTTACAAAGACGTTT 2877

Qy  1381 GGATAACATGAACTTCAAGTGAGTGAAGTGAGGCGAAAAAGTCTCTCAACATTAGGTCCCA 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2878 GGATAACATGAACTTCAAGTGAGTGAAGTGAGGCGAAAAAGTCTCTCAACATTAGGTCCCA 2937

Qy  1441 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGACCGTCTGCAGGAACTTCTGGT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2938 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCCTTTCTGCAGGAACTTCTGGT 2997

Qy  1501 GTGGCTACAGCTGAAAGATGAAGATGAATTAAGCCGGCACCCATTGGAGGCGACTTTC 1560
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2998 GTGGCTACAGCTGAAAGATGAAGATGAATTAAGCCGGCACCCATTGGAGGCGACTTTC 3057

Qy  1561 AGCAGTTCAGAGAGCAGAGAACTGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3058 AGCAGTTCAGAGAGCAGAGAACTCACAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGA 3117

Qy  1621 ACCTGTAATCATGAGTACTCTTGAGACTCTGTACGAATATATTTCGACAGAGCAGCCTTTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3118 ACCTGTAATCATGAGTACTCTTGAGACTGTGTACGAATATTTTCGACAGAGCAGCCCTTGA 3177

Qy  1681 AGGACTAGAGAAACTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGGAGGAGAGGCCCAGAA 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3178 AGGACTAGAGAAACTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGGAGGAGAGGCCCAGAGA 3237

Qy  1741 TGTCACTCGGCTTCTACCGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3238 TGTCACTCGGCTTCTACCGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 3297

Qy  1801 CCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGGAACTTGAAAGACTCCAGGAACT 1860
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3298 CCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGGAACTTGAAAGACTCCAGGAACT 3357

Qy  1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGCTGATCAAGGGATC 1920
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Best Local Similarity  84.6%;   Pred. No.  0;   Indels  324;   Gaps  1;
Matches 1777;  Conservative  0;  Mismatches  0;
```

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCCTGACCCTACACGAGCCCATT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCCTGACCCTACACGAGCCCATT  959

Db   1678 ------------------------------------------------------------
Qy   1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGACCTGGA 1140

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGAG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGAG 1019

Db   1678 ------------------------------------ACTCATAGATTACTGCAACAGTTCCCCTGACCTGGA 1715
Qy   1141 AAAGTTTCTTGCCTGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACACAGGATGCTAC 1200

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAGTATTATCGTGCTTCTTTC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAGTATTATCGTGCTTCTTTC 1079

Db   1716 AAAGTTTCTTGCCTGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1775
Qy   1201 CCGTAAGGAAAGGTCCTCAAGGACTCCAAGGAGTAAAGAGCTGATGAAACAATGGCA 1260

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 1139

Db   1776 CCGTAAGGAAAGGTCCTCAAGGACTCCAAGGAGTAAAGAGCTGATGAAACAATGGCA 1835
Qy   1261 AGACCTCCAAGTGAAATGAAGTCACACAGATGTTATCACAACCTGATGAAAACAG 1320

Qy    241 CCAGTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGGCCGGGTTGG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGGCCGGGTTGG 1199

Db   1836 AGACCTCCAAGTGAAATTGAAGTCACACAGATGTTATCACAAAGACGTTT 1895
Qy   1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCAGTCCCA 1380

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA 1259

Db   1896 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCAGTCCCA 1955
Qy   1381 GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGGTCCCA 1440

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Db   1956 GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAGAACTTCGGT 2015
Qy   1441 TTTGAAGCCAGTTCTGACCAGTGGAAGGTCTGCGAAGGTGAAGTGTCAGGAACTTCTGGT 1500

Qy    421 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1379

Db   2016 TTTGAAGCCAGTTCTGACCAGTGGAAGGTCTGCGAAGGTGAAGTGTCAGGAACTTCTGGT 2075
Qy   1501 GTGGCTACAGCTGAAAGATGATGAAGATGAATTAAGCCGGCAGCCACCTATTGGAGGCGACTTTCC 1560

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAAGAACAAGGAAAATGAGAGGAAGA  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAAGAACAAGGAAAATGAGAGGAAGA 1439

Db   2076 GTGGCTACAGCTGAAAGATGATGAAGATGAATTAAGCCGGCAGCCACCTATTGGAGGCGACTTTCC 2135
Qy   1561 AGCAGTTCAGAAGCAGAAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620

Qy    541 GCCTCCTTGAGACCTGATCTTGAAGACTTCAGGGTGCAACTGCTCAACTTCTCACTCACAGTGGTGTGTAGT  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 GCCTCCTTGAGACCTGATCTTGAAGACTACAGTGGTGTGTAGT 1499

Db   2136 AGCAGTTCAGAAGCAGAAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 2195
Qy   1621 ACCTGTAATCATGATGACTCTTGAGACTGTACAGAGCAGCCTTTTGA 1680

Qy    601 AGAAGATCTAGAACAAGAGAACAAGTCAGGGTGCTCAATTCTCTCACTCACATGGTGTGGTAGT  660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 AGAAGATCTAGAACAAGAGAACAAGTCAGGGTGCTCAATTCTCTCACTCACATGGTGTGGTAGT 1559

Db   2196 ACCTGTAATCATGATGACTCTTGAGACTGTACAGAGCAGCCTTTTCA 2255
Qy   1681 AGGACTAGAGAAACTCTACCAGAGGCTGCCTCCTGAGAGAGCCCAGAA 1740

Qy    661 TGATGAATCTAGTGAGGACCGCAAACATCGTAGATGGACAGAAGACCGTGGGTTCTTTTACAAGACAT  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1560 TGATGAATCTAGTGAGGACCGCAAACATCGTAGATGGACAGAAGACCGTGGGTTCTTTTACAAGACAT 1619

Db   2256 AGGACTAGAGAAACTCTACCAGAGGCTGCCTCCTGAGAGAGCCCAGAA 2315
Qy   1741 TGTCACTCGGCTTCTTCACGAAAGGCAGGGTGAGAGGAGTGGGAAAATTGAA 1800

Qy    721 AGATCGATGGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC-- 1677
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1620 AGATCGATGGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC  1677

Db   2316 TGTCACTCGGCTTCTTCACGAAAGGCAGGGTGAGAGGAGTGGGAAAATTGAA 2375
Qy   1801 CCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1860

Qy    781 CCTTCTCAAATGGCAACGTGCAACGTCTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA  840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 ------------------------------------------------------------ 1677

Db   2376 CCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 2435
Qy   1861 TCAAGAGGGCCACCGGATGAGCTGGACCTCAAGTGAGGTGATCAAGGGATC 1920

Qy    841 AAAAGAAGATGCAGTGAACAAGAGATTCACCACAACCTGGCTTTAAAGATCAAAATGAAATGTT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 ------------------------------------------------------------ 1677

Db   2436 TCAAGAGGGCCACCGGATGAGCTGGACCTCAAGTGAGGTGATCAAGGGATC 2495
Qy   1921 CTGGCAGCCGGTGGGCGATCTCCTCATTGACTCCTCCAAGATCACCTCGAGAAGTCAA 1980

Qy    901 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCCGGATCTAGAAAAGAGAAAAAGCAATCCAT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 ------------------------------------------------------------ 1677

Db   2496 CTGGCAGCCGGTGGGCGATCTCCTCATTGACTCCTCCAAGATCACCTCGAGAAAGTCAA 2555
Qy   1981 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAAGAACGTGAGCCCACGTCAATGACCTTGC 2040

Qy    961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 ------------------------------------------------------------ 1677

Db   2556 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAAGAACGTGAGCCCACGTCAATGACCTTGC 2615
Qy   2041 TCGGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGAAGA 2100

Qy   1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 ------------------------------------------------------------ 1677

Db   2616 TCGGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2675

Qy   2101 C 2101
Db   2676 C 2676
```

RESULT 7

AD37257

AD AD37257 standard; DNA; 4825 BP.

XX

ID AAD37257;

XX

XX 21-AUG-2002 (first entry)

XX

DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.

XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX

OS Homo sapiens.
OS Unidentified.
OS Chimeric.

XX

CX WO200183695-A2.

XX

PD 08-NOV-2001.

XX

PF 27-APR-2001; 2001WO-US013677.

XX

PR 28-APR-2000; 2000US-0200777P.

XX

PA (XIAO/) XIAO X.

XX

PI Xiao X;

XX

PR WPI; 2002-049342/06.

XX

PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.

XX

PS Example 1; Page 61-62; 71pp; English.

XX

CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence

XX

SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;

Query Match   68.7%; Score 1443; DB 6; Length 4825;
Best Local Similarity   84.6%; Pred. No. 0;
Matches 1777   Conservative 0; Mismatches 0; Indels 324; Gaps 1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1716

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1776

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy   241 CCAGTTCATACTCATGAGGGTACATGATTGATTGACAGCCATCAGGGCCGGGTTGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTCATACTCATGAGGGTACATGATTGATTGACAGCCATCAGGGCCGGGTTGG 1956

Qy   301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATATCAGAAGATGAAGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATATCAGAAGATGAAGA 2016

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
            |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2136

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA 2196

Qy   541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256

Qy   601 AGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 AGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 2316

Qy   661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
            |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376

Qy   721 AGATGCGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAT 780
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGAC-- 2434

Qy   781 CCTTCTCAAATGCCAACGTCTTACTGGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2435   2434

Qy   841 AAAAGAGAGATGCAGTGAACAGATTCACACAACTGGCTTTAAAGATCAAATGAAAATGTT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2435   2434

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2435   2434

Qy   961 GGGCAAACTGTATTCACTCAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2435   2434

Qy   1021 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2435   2434

Qy   1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACGCAACAGTTCCCCTGACCTGGA 1140
            ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   2435   ACTCATAGATTACGCAACAGTTCCCCCTGACCTGGA 2472

Qy   1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATCTCCTACAGGATGCTAC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2473 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATCTCCTACAGGATGCTAC 2532

Qy   1201 CCGTAAGGAAAGGCCTCCAAGGAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCA 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2533 CCGTAAGGAAAGGCCTCCAAGGAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCA 2592

Qy   1261 AGACCTCCACGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAAACAG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2261 AGAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 2316

OS Homo sapiens.
OS Cytomegalovirus.
OS Unidentified.
XX
PN WO200183695-A2.
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 68-70; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;

Query Match   68.7%; Score 1443; DB 6; Length 4848;
Best Local Similarity   84.6%;   Pred. No. 0;
Matches 1777; Conservative   0; Mismatches   0; Indels 324; Gaps   1;

```
Qy          1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCTACGGAGCCCATT   60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCTACACAGCCCATT 1739

Qy         61 TCCTTCACAGCATTGGAAGTCTCGAAGACAAGTCATTGGCAGTTCATTGGATGAGAG  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 TCCTTCACAGCATTGGAAGTCTCGAAGACAAGTCATTGGCAGTTCATTGGATGAGAG 1799

Qy        121 TGAAGTAAACCTGGACCGTATCAAGCACCAACAGTTTAGAAGAGATTATCGTGGCTTCTTTC  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 TGAAGTAAACCTGGACCGTATCAAGCACCAACAGTTTAGAAGAGATTATCGTGGCTTCTTTC 1859

Qy        181 TGCTGAGGACACATTGCAAGCACAAGCACACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1860 TGCTGAGGACACATTGCAAGCACAAGCACACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1919

Qy        241 CCAGTTTCATACTCATCATGAGGGGTACACAGGGGTACATGAGGGCAAAATTATCAGAAGATGAAGA  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCAGTTTCATACTCATCATGAGGGGTACACAGGGGTACATGAGGGCAAAATTATCAGAAGATGAAGA 1979

Qy        301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2039

Qy        361 AACTGAAGTACAAGAGCAAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2040 AACTGAAGTACAAGAGCAAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2099

Qb         2593 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTATCACAACCTGGATGAAAACAG 2652
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy         1321 CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCGTTACAAAGACGTT 1380

Qb         2653 CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCGTTACAAAGACGTT 2712

Qy         1381 GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTTCTCAACATTAGGTCCCA 1440
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         2713 GGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 2772

Qy         1441 TTTGGAAGCCAGTTCTGACCAGTGAAGGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 1500
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         2773 TTTGGAAGCCAGTTCTGACCAGTGAAGGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGT 2832

Qy         1501 GTGGCTACAGCTGAAAGATGATGAAGATGAATTAAGCCGGCAGGCACCTATTGAGGGCGACTTTCC 1560
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         2833 GTGGCTACAGCTGAAAGATGATGAAGATGAATTAAGCCGGCAGGCACCTATTGAGGGCGACTTTCC 2892

Qy         1561 AGCAGTTCAGAGAGCAGAACAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         2893 AGCAGTTCAGAAGCAGAACAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 2952

Qy         1621 ACCTGTAATCATGAGTACTCTTGACACTGTACGAATAATTCTGACAGAGCAGCCTTTTGA 1680
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         2953 ACCTGTAATCATGAGTACTCTTGACACTGTACGAATAATTCTGACAGAGCAGCCTTTTGA 3012

Qy         1681 AGGACTAGAGAAACTTCTACCAGGAGAGCTGCCTCCTGGAGGAGAGGCCCAGAA 1740
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         3013 AGGACTAGAGAAACTTCTACCAGGAGAGCTGCCTCCTGGAGGAGAGAGGCCCAGAA 3072

Qy         1741 TGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 1800
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         3073 TGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 3132

Qy         1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1860
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         3133 CCTGCACTCCCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 3192

Qy         1861 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 1920
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         3193 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 3252

Qy         1921 CTGGCAGCCCCGTGGGCGGATCTCCTCATTGACTCTCCAAGATCACCTCGAGAAAGTCAA 1980
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         3253 CTGGCAGCCCCGTGGGCGGATCTCCTCATTGACTCTCCAAGATCACCTCGAGAAAGTCAA 3312

Qy         1981 GGCACTTCGAGGAGAGAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qb         3313 GGCACTTCGAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3372

Qy         2041 TCGCCGACGTTACCACTTGGGCATTCAGCTCTCAGCACTCTGGAGAA 2100
                   ||||||||||||||||||||||||||||||||||||||||||||||||
Qb         3373 TCGCCCAGCTTACCACTTGGGCATTCAGCTCTCAGCACTCTGGAGAA 3432

Qy         2101 C 2101
                   |
Qb         3433 C 3433
```

RESULT 8
AD37263
D AAD37263 standard; DNA; 4848 BP.
X
C AAD37263;
X
T 21-AUG-2002 (first entry)
X
E Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
X
W Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.

```
Qy    421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAATGATCTTCCAGAATCAGAAACT  480
          |||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   2100 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTTCCAGAATCAGAAACT 2159

Qy    481 GAAAGAGTTGAATGACTGGCTACAAAAACAGAGAGAAAGACAAGGAAAGTGGAGGAAGA  540
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   2160 GAAAGAGTTGAATGACTGGCTACAACAACAGAGAGAAAGACAAGGAAAGTGGAGGAAGA 2219

Qy    541 GCCTCTTGGACCTGATCTTGAAGACTACAACATCCAAGTACAACACACAACATAGGTGCTTCA  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2220 GCCTCTTGGACCTGATCTTGAAGACTACAACATCCAAGTACAACACACAACATAGGTGCTTCA 2279

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2280 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT 2339

Qy    661 TGATGAATCTAGTGGAGATCACGCAACTGCTCGTTGGAAGAACAACTTAAGGTATTGGG  720
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   2340 TGATGAATCTAGTGGAGATCACGCAACTGCTCGTTGGAAGAACAACTTAAGGTATTGGG 2399

Qy    721 AGATCGATGGGCAAACATCTGTAGATGGACAAGAACCGGTTGGGTTCTTTTACAAGACAT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2400 AGATCGATGGGCAAACATCTGTAGATGGACAAGAACCGGTTGGGTTCTTTTACAAGAC- 2457

Qy    781 CCTTCTCTCAAATGGCAACGTCTTACTGAAGACGAGTGCCTTTTAGTGCATGGCTTTCAGA  840
          -----------------------------------------------------------
Db   2458 ------------------------------------------------------------

Qy    841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT  900
          -----------------------------------------------------------
Db   2458 ------------------------------------------------------------

Qy    901 ATCAAGTCTTCAAAAACTGGCGCGTTTAAAAGCGGATCCTAGAAAAGAAAAGAAAGCAATCCAT  960
          -----------------------------------------------------------
Db   2458 ------------------------------------------------------------

Qy    961 GGGCAAACTGTATTCACTCAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
          -----------------------------------------------------------
Db   2458 ------------------------------------------------------------

Qy   1021 CCAGAAGACGGGAGCATGGCTGGATAACTTTGCCCGTGTGTGGGATAATTTAGTCCAAAA 1080
          -----------------------------------------------------------
Db   2458 ------------------------------------------------------------

Qy   1081 ACTTGAAAAGAGTACAGCACAGACTACAGCAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
          --ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2495
Db   2458 --ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2495

Qy   1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1200
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   2496 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 2555

Qy   1201 CCGTAAGGAAAAGGCTCCTAGAGAGACTCCAAGGGGAGTAAAAGCTGATGAAACAATGGCA 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2556 CCGTAAGGAAAAGGCTCCTAGAGAGACTCCAAGGGGAGTAAAAGCTGATGAAACAATGGCA 2615

Qy   1261 AGACCTCCAAGGTGAAATTGAAGCTGAAGCTGAAGTGTTTATCACAACCTGGATGAAAACAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2616 AGACCTCCAAGGTGAAATTGAAGCTGAAGCTGAAGTGTTTATCACAACCTGGATGAAAACAG 2675

Qy   1321 CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAAGACGTTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2676 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAAGACGTTT 2735

Qy   1381 GGATAACATGAACTTCAGTGACTTCAGTGAAGCGTCTGACACCGGAAGCGTCTGACACCTTCTGGT 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2736 GGATAACATGAACTTCAGTGACTTCAGTGAAGCGTCTGACACCGGAAGCGTCTGACACCTTCTGGT 2795

Qy   1441 TTTCGAAGCCAGTTCTGACCAGTGGAGCGTCTGCCACCTTTCTCTGCAGGAACTTCTGGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2796 TTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCCACCTTTCTCTGCAGGAACTTCTGGT 2855

Qy   1501 GTGGCTACAGCTGAAAGATGAATGATGAAATTAAGCCGGCAGGCCACCTATTGGAGGCGACTTTCC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2856 GTGGCTACAGCTGAAAGATGAATGATGAATTAAGCCGGCAGGCCACCTATTGGAGGCGACTTTCC 2915

Qy   1561 AGCAGTTCAGAGAGCAGAACGAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2916 AGCAGTTCAGAGAGCAGAACGAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 2975

Qy   1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGACCCTTGGA 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2976 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGACCCTTGGA 3035

Qy   1681 AGGACTAGAGAAACTTCTACCAGGAGCCCAGAGAGTGCTCCTGAGGAGGAGAGCCCAGAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3036 AGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTGCCCTCCTGAGGAGGAGAGCCCCAGAA 3095

Qy   1741 TGTCACTCGGCTTCTACGAAGCAGGTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3096 TGTCACTCGGCTTCTACGAAGCAGGTGAGGAGTCAATACTGAGTGGGAAAAATTGAA 3155

Qy   1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAAACT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3156 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAAACT 3215

Qy   1861 TCAAGAGGCCACGGATGAGTCTGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3216 TCAAGAGGCCACGGATGAGTCTGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATC 3275

Qy   1921 CTGGCAGCCGCGTGGGCGCCGATCTCCATTGACTCTCCCAAGATCACCTCGAAGAGTCAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3276 CTGGCAGCCGCGTGGGCGCCGATCTCCATTGACTCTCCCAAGATCACCTCGAGAAGTCAA 3335

Qy   1981 GGCACTTCGAGGAGAATTGCGGCCGCACTGAGCCACGTCAATGACCTTGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3336 GGCACTTCGAGGAGAAATTGCGCCCTCTGAAAGAGACAGCGAGCCACGTCAATGACCTTGC 3395

Qy   2041 TCGCCAGTTACCACTTGGGCATTCAGCCTGTAAACCTCTCAGCACTCTGAAGA 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3396 TCGCCAGTTACCACTTTGGGCATTCAGCCTGTAAACCTCTCAGCACTCTGAAGA 3455

Qy   2101 C 2101

Db   3456 C 3456
```

Xiao X;

WPI; 2002-049342/06.

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

Example 1; Page 70-71; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a small polyA signal sequence

Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

Query Match   68.7%;   Score 1443;   DB 6;   Length 5060;
Best Local Similarity   84.6%;   Pred. No. 0;
Matches 1777;   Conservative 0;   Mismatches 0;   Indels 324;   Gaps 1;

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTATGTCAAGAACTCTGACCCTACACGGAGCCCATT   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1892 GAGCTATGCCTACACACAGGCTGCTATGTCAAGAACTCTGACCCTACACGGAGCCCATT 1951

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1952 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG 2011

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2012 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 2071

Qy    181 TGCTGAGGACACAATTGCAAGCACAAGGACACAAGGAGAGATTTCTAATGATGTGAAGTG  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2072 TGCTGAGGACACAATTGCAAGCACAAGGACACAAGGAGAGATTTCTAATGATGTGGAAGTG 2131

Qy    241 CCAGTTCATACTCATCATGAGGGGTACACTGACGATTTGACAGCCCCATCAGGGCGGGTTGG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2132 CCAGTTCATACTCATGAGGGGTACACTGACGATTTGACAGCCCATCAGGGGCGGGTTGG 2191

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGAACAGGAAATTATCAGAAGATGAAGA  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2192 TAATATTCTACAATTGGGAAGTAAGCTGATTGAACAGGAAATTATCAGAAGATGAAGA 2251

Qy    361 AACTGAAGTACAAGAGCAGAAGGATGAATCTCCTAAATTCAGAAGATGGGAATGCCTCAGGGTAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2252 AACTGAAGTACAAGAGCAGAAGGATGAATCTCCTAAATTCAGAAGATGGGAATGCCTCAGGGTAGC 2311

Qy    421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2371

Qy    481 GAAAGAGTTGATGACTGGCTAACAAAACAGAGAAGAACAAGGAAAATGAGAGAAGA  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2372 GAAAGAGTTGATGACTGGCTAACAAAACAGAGAAGAACAAGGAAAATGAGAGAAGA 2431

Qy    541 GCCTCTTGGACCTGATCTTGAAGACTCTGAAGACAGTCAGGGTCAATTCTGAAGACAAGCAAGGTGCTTCA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2432 GCCTCTTGGACCTGATCTTGAAGACTCTGAAGACTAACAACAACAGCCAAGTAGACTACAACAACATAAGGTGCTTCA 2491

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGTGGTAGT  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2492 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGTGGTAGT 2551

Qy    661 TGATGAATCAGTGGAGATCACGCAACTCGCTGCTT...TTGGAAGAACAACTAAGGTATTGGG  720
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2552 TGATGAATCAGTGGAGATCACGCAACTCGCTGCTT...TTGGAAGAACAACTAAGGTATTGGG 2611

Qy    721 AGATCGATGGGCAAACATCTGTAGATGACAGAGAACCGCTGGGTTCTTTTACAAGACAT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2612 AGATCGATGGGCAAACATCTGTAGATGACAGAGAACCGCTGGGTTCTTTTACAAGAC-- 2669

Qy    781 CCTTCCAAATGGCACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  840
Db   2670 ------------------------------------------------------ 2669

Qy    841 AAAAGAAGATGCAGTGAACAAGATTCACCACACTGGCTTAAAGATCAAAATGAAATGTT  900
Db   2670 -------------------------------------------------------- 2669

Qy    901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAGCAGAAAAAGCAATCCAT  960
Db   2670 ------------------------------------------------------------ 2669

Qy    961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
Db   2670 -------------------------------------------------------- 2669

Qy   1021 CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTAGTCCAAAA 1080
Db   2670 ------------------------------------------------------- 2669

Qy   1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTCCCCTGACCTGA 1140
          --------------||||||||||||||||||||||||||||||||||||||||
Db   2670 ----------ACTCATAGATTACTGCAACAGTCCCCTGACCTGA 2707

Qy   1141 AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCTTACAGAGATGCTAC 1200
          --------||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2708 -----AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCCAATGTCTTACAGAGATGCTAC 2767

Qy   1201 CCGTAAGGAAAGGCTTCTCCAAGGAGAGACTCTCCAAGGAGAGTAAAAGAGAGCTGATGAAACAATGGCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2768 CCGTAAGGAAAGGCTTCTCCAAGGAGAGACTCTCCAAGGAGAGTAAAAGAGAGCTGATGAAACAATGGCA 2827

Qy   1261 AGACCTCCAAGTGAAATTGAAGTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2828 AGACCCTCCAAGTGAAATTGAAGTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2887

Qy   1321 CCAAAAAATCCTGACATGATGAAAGGTTCCGATGATCCAGTCCGTGTTACAAAGACGTTT 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2888 CCAAAAAATCCTGACATGATGAAAGGTTCCGATGATCCAGTCCGTGTTACAAAGACGTTT 2947

Qy   1381 GGATAACATGACAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2948 GGATAACATGACAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3007

Qy   1441 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCCACCTTTCTCTCGCAGGGAACTTCTGGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3008 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCCACCTTTCTCTCGCAGGGAACTTCTGT 3067

Qy   1501 GTGGCTACAGCTGAAAGTGAAAGATGATGAATGAATTAAGCCCGGCACCTATTGGAGGGGACTTTCC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3068 GTGGCTACAGCTGAAAGTGAAAGATGATGAATGAATTAAGCCCGGCACCTATTGGAGGGGCGACTTTCC 3127

Qy   1561 AGCAGTTCAGAGAGCAGAAGACGATGACTATGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3128 AGCAGTTCAGAGAGCAGAAGACGATGACTATGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 3187

Qy   1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACAGATATTTCTGACAGAGTATTCTGACAGCCTTTGGA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3188 ACCTGTAATCATGAGTACTCTTGAGACTGTACAGATATTTCTGACAGAGTATTCTGACAGCCTTTGGA 3247

Qy   1681 AGGACTAGAAGAAACTCTACCAGAGCGCCCAGAGAGCGCTGCCTCCTGAGGAGAGAGCGCCCAGAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3248 AGGACTAGAGAAACTCTACCAGAGCGCCCAGAGAGCGCTGCCTCCTGAGGAGAGAGCGCCCAGAA 3307
```

CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 U; 0 Other;

Query Match   61.5%; Score 1293; DB 6; Length 8689;
Best Local Similarity   72.5%; Pred. No. 0;
Matches 2101; Conservative   0; Mismatches   0; Indels 798; Gaps   1;

```
Qy         1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCGGCCATT 60
             |||||||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||
Db      1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCGGCCATT 1158

Qy        61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCATTGGCAGTTCATTGATGATGGAGAG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCATTGGCAGTTCATTGATGATGGAGAG 1218

Qy       121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGAGTATTATCGTGGCTTCTTTC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGAGTATTATCGTGGCTTCTTTC 1278

Qy       181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy       241 CCAGTTCATCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1339 CCAGTTCATCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy       301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACCAGAGAAATTATCAGAAGATGAAGA 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA 1458

Qy       361 AACTGAAGTACAAGAGCAGATGAATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGTAGC 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1459 AACTGAAGTACAAGAGCAGATGAATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGTAGC 1518

Qy       421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1519 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy       481 GAAAAGAGTTGAATGACTGGCTAACAAAAACAAGGAAATGGAGGAAGA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1579 GAAAAGAGTTGAATGACTGGCTAACAAAAACAAGGAAATGGAGGAAGA 1638

Qy       541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 1698

Qy       601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATCTCTCCACTCACATGTGGTGGTAGT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATCTCTCCACTCACATGTGGTGGTAGT 1758

Qy       661 TGATGAATCTAGTGGAGATCAGCGAACTACGCAACTTAGGTATTGGG 720
             |||||||||||||||||||||||||||||||||||||||||||||||
Db      1759 TGATGAATCTAGTGGAGATCAGCGAACTACGCAACTTAGGTATTGGG 1818

Qy       721 AGATCGATGGGCAAACATCTGTAGATGGAACATCTGTAGATGGAACACCGCTGGGTTCTTTTACAAGACAT 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1819 AGATCGATGGGCAAACATCTGTAGATGGAACATCTGTAGATGGAACACCGCTGGGTTCTTTTACAAGACAT 1878

Qy       781 CCTTCTCAAATGGCAACGTCTACTGAAGAACAGTGCCTTTAGTCATGGCTTTCAGA 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1879 CCTTCTCAAATGGCAACGTCTACTGAAGAACAGTGCCTTTAGTCATGGCTTTCAGA 1938

Qy       841 AAAAGAGATGCAGTGAACAAGATCCACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1939 AAAAGAGATGCAGTGAACAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998

Qy       901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTGAAAAGAAAAAGCAATCCAT 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTGAAAAGAAAAAGCAATCCAT 2058
```

```
Qy      1741 TGTCACTCGGCCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3308 TGTCACTCGGCCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3367

Qy      1801 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATAGACCCTTGAAAGACTCCAGGAACT 1860
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3368 CCTGCACTCCGCTGACTGGCGAGAGAAAAATAGATAGACCCTTGAAAGACTCCAGGAACT 3427

Qy      1861 TCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3428 TCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATC 3487

Qy      1921 CTGGCAGCCCGTGGGCCGATCTCCATTGACTCTTCCAAGATCACCTGAGAAGTCAA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3488 CTGGCAGCCCGTGGGCCGATCTCCATTGACTCTTCCAAGATCACCTGAGAAGTCAA 3547

Qy      1981 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAATGACCTTGC 2040
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3548 GGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAATGACCTTGC 3607

Qy      2041 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3608 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3667

Qy      2101 C 2101
             |
Db      3668 C 3668
```

RESULT 10
ABK82000
ID   ABK82000 standard; DNA; 8689 BP.
XX
AC   ABK82000;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaH2-R19.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO20022056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
PR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 15; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition

```
  961  GGGCAAACTGTATTCACTCAAACAGAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  1020
       |||||||||||||||||||||||||||||||||||
 2059  GGGCAAACTGTATTCACTCAAACAAGAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  2118

 1021  CCAGAAGACGGAAGCATGGCTGGATAACCTTTGCCCGGTGTTGGGATAATTAGTCCAAAA  1080
       |||||||||||||||||||||||||
 2119  CCAGAAGACGGAAGCATGGCTGGATAACCTTTGCCCGGTGTTGGGATAATTAGTCCAAAA  2178

 1081  ACTTGAAAAGAGTACAGCACAG-------------------------------------  1102
       ||||||||||||||
 2179  ACTTGAAAAGAGTACAGCACAGATTTCACAGCAGCCTAGCTCCTGACTGACCAC       2238

 1103  ---------------------------------------------------------  1103
 2239  TATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAACCTGGTTACTAAGGA   2298

 1103  ---------------------------------------------------------  1103
 2299  AACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTGGACGTACCTGCTCTGGC  2358

 1103  ---------------------------------------------------------  1103
 2359  AGATTTCAACCGGGCTTGGACAGAACTTACCGACTGGCTTTCTCTGCTTGATCAAGTTAT  2418

 1103  ---------------------------------------------------------  1103
 2419  AAAATCACAGAGGGTGATGGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA   2478

 1103  ---------------------------------------------------------  1103
 2479  GAAGGCAACAATGCAGGATTTGGAACAGAGGCGTCCCCAGTTGGGAAGAACTCATTACCGC  2538

 1103  ---------------------------------------------------------  1103
 2539  TGCCCAAAATTTGAAAAACAAGACCAGCAATCAAGAGAACAATCATTACCGGATCG      2598

 1103  ---------------------------------------------------------  1103
 2599  AATTGAAAGAATTCAGAATCAGTGGATGAAGTACAGACAGAAACCTTCAGAACCGGAGGCA  2658

 1103  ---------------------------------------------------------  1103
 2659  ACAGTTGAATGAATGTTAAAGGATTCAACACAATGGCTGGAAGTCTGGAAGTCTGGAAGAAGCTGA  2718

 1103  ---------------------------------------------------------  1103
 2719  GCAGGTCTTAGGACAGGCCAGAGCCAAGCTTGAGTCATGCGGAAGCAGAGGGTCCCTATACAGT  2778

 1103  ---------------------------------------------------------  1103
 2779  AGATGCAATCAAAAGAAGAAAATCACAGAAACCAGCAGTGGCCAAAGACCTCCGCCAGTG  2838

 1103  ---------------------------------------------------------  1103
 2839  GCAGACAAATGTGTAGTGTGGCAAATGACTGGCCCTGAAAACTTCTCCGGGATTATTCTGC  2898

 1103  ------------------------------------ACTCATAGATTACTGCAACA   1122
                                          |||||||||||||||||||||
 2959  TCATAAAGGGGTGAGTGAGCCGAGCGGCTGCTGTTTGTGAGGAAAACTCATAGATTACTGCAACA  3018

 1123  GTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGCCTGCTTACAGAGAGCTAAACAACTGCCAA   1182
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 3019  GTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGCCTGCTTACAGAGAGCTAAACAACTGCCAA   3078

 1183  TGTCCTACAGGATGCTACCCGTAAGGAAAGGGCTTCCTAGGAAGAGACTTCCAAGGAGTAAAAGA   1242
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 3079  TGTCCTACAGGATGCTACCCGTAAGGAAAGGGCTTCCTAGGAAGAGACTTCCAAGGAGTAAAAGA   3138
```

```
 Qy   1243  GCTGATGAAACAATGGCAGAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCA   1302
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3139  GCTGATGAAACAATGGCAAGAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCA   3198

 Qy   1303  CAACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGAGATCCTGGAAGGTTCCGATGATGCAGT   1362
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3199  CAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGAGATCCTGGAAGGTTCCGATGATGCAGT   3258

 Qy   1363  CCGTACAAAGACGTTTGGATAACATGAACTTCAAGTGGACTGAACTTCGGAAAAAGTC   1422
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3259  CCGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGACTGAACTTCGGAAAAAGTC   3318

 Qy   1423  TCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGGCTCTGACCTTTC   1482
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3319  TCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGGCTCTGACCTTTC   3378

 Qy   1483  TCTGCAGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC   1542
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3379  TCTGCAGGAACTTCTGGTGTGGCTACACTGAAAGATGATGAATTAAGCCGGCAGGCACC   3438

 Qy   1543  TATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGGAACGATGTACATAGGGCCTTCAAGAG   1602
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3439  TATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGGAACGATGTACATAGGGCCTTCAAGAG   3498

 Qy   1603  GGAATTGAAAACTAAAGAACCTGTAATCATGAGAACCTTCTTGAGACTGTTACGAATATTTCT   1662
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3499  GGAATTGAAAACTAAAGAACCTGTAATCATGAGAACCTTCTTGAGACTGTTACGAATATTTCT   3558

 Qy   1663  GACAGAGCAGCGCCTTGGAAGAACTAGAGAACTCTACCAGGAAAGCAGGGCTGAGGAGTCGCTCC   1722
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3559  GACAGAGCGCCTTTGGAAGAACTAGAGAACTCTACCAGGAAAGCAGGGCTGAGGAGTCGCTCC   3618

 Qy   1723  TGAGGAGAGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTCGAAAGATGAGACCCT   1782
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3619  TGAGGAGAGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTCGAAAGATGAGACCCT   3678

 Qy   1783  TGAGTGGAAAAATTGAACCTGACTGGCAGAGGAAATAGATGAGACCCT   1842
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3679  TGAGTGGGAAAAATTGAACCTGACTGGCAGAGGAGAAAATAGATGAGACCCT   3738

 Qy   1843  TGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGC   1902
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3739  TGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGC   3798

 Qy   1903  TGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCCTCTCCAAGA   1962
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3799  TGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCCTCTCCAAGA   3858

 Qy   1963  TCACCTGAGAAGTCAGGCACTTGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAG   2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3859  TCACCTGAGAAGTCAGGCACTTGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAG   3918

 Qy   2023  CCAGTCAATGACCTTGCTCGCCGCCAGCTTACACTTTGGGCATTCAGCTCTCACCGTATAA   2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db   3919  CCAGTCAATGACCTTGCTCGCCGCCAGCTTACACTTTGGGCATTCAGCTCTCACCGTATAA   3978

 Qy   2083  CCTCAGCACTCTGAAAGAC   2101
            |||||||||||||||||||
 Db   3979  CCTCAGCACTCTGAAAGAC   3997
```

XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   misc_recomb   2847..2848
FT   /*tag= a
FT   /label= S4_junction_site
FT   /note= "dystrophin gene is split at this site and
FT   attached to an intein sequence. The resulting fragment is
FT   used to produce plasmid pSP4"
FT   misc_recomb   2952..2953
FT   /*tag= b
FT   /label= S3_junction_site
FT   /note= "dystrophin gene is split at this site and
FT   attached to an intein sequence. The resulting fragment is
FT   used to produce plasmid pSD3"
FT   misc_recomb   3198..3199
FT   /*tag= c
FT   /label= S2_junction_site
FT   /note= "dystrophin gene is split at this site and
FT   attached to an intein sequence. The resulting fragment is
FT   used to produce plasmid pSD2"
FT   misc_recomb   3300..3301
FT   /*tag= d
FT   /label= S1_junction_site
FT   /note= "dystrophin gene is split at this site and
FT   attached to an intein sequence. The resulting fragment is
FT   used to produce plasmid pSD1"
XX
XN   WO20129243-A1.
XX
XD   26-APR-2001.
XX
XF   13-OCT-2000; 2000WO-CA001216.
XX
XR   15-OCT-1999; 99US-0159868P.
XX
PA   (UYDA-) UNIV DALHOUSIE.
PA   (UYPI-) UNIV PITTSBURGH.
XX
PI   Paul XL, Xiao X;
XX
XR   WPI; 2001-367297/38.
XX
FT   Use of spontaneous or automatic protein splicing to join two or more
FT   peptides at junction site involves expressing extein peptides having co-
FT   reacting portions of split intein attached to them, so that peptides
FT   splice.
XX
XS   Example 1; Fig 2; 81pp; English.
XX
XC   The invention relates to a method directed to the use of spontaneous or
XC   automatic protein trans-splicing to join two or more peptides having co-
XC   automatic protein trans-splicing to join two or more peptides at junction
XC   site. This method involves expressing extein peptides having co-reacting
XC   portions of split intein attached to them, so that peptides will splice
XC   automatically under suitable conditions. The invention also provides
XC   methods for circumventing virion packaging size limitations in
XC   recombinant virus particle, by splitting a coding region for a protein to
XC   be delivered into two or more extein genes, which are packaged in
XC   separate virus particles and are co-delivered in a target cell for the
XC   expression and for subsequent trans-splicing to form the complete
XC   protein. In particular, the method is used for trans-splicing human
XC   dystrophin and in gene therapies of recombinant adeno-associated virus
XC   (AAV) particles that encode trans-spliced dystrophin, for treating
XC   diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular
XC   dystrophy (BMD). The present sequence is human dystrophin gene (Becker
XC   form).
XX
XQ   Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 U; 0 Other;

Query Match   53.8%; Score 1131; DB 5; Length 5952;
Best Local Similarity 68.6%; Pred. No. 1.1e-306;
Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACGCGGCCCATT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   891 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACGCGGCCCATT 950

Qy   61 TCCTTCACAGCATTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGGAGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   951 TCCTTCACAGCATTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGGAGAG 1010

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAGAAGTATTATCGTGGCTTCTTTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1011 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAGAAGTATTATCGTGGCTTCTTTC 1070

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1130

Qy   241 CCAGTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1131 CCAGTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG 1190

Qy   301 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACACAGGAAAATTATCAAGATGAAGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1191 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACACAGGAAAATTATCAAGATGAAGA 1250

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1251 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC 1310

Qy   421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1370

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAACAGAGAAAGAACAAGGAAAATGGGAGGAAGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1371 GAAAGAGTTGAATGACTGGCTAACAAACAGAGAAAGAACAAGGAAAATGGGAGGAAGA 1430

Qy   541 GCCTCTTGGACCTGATCTTGAAGACATTCTGAAGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 GCCTCTTGGACCTGATCTTGAAGACATTCTGAAGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1490

Qy   601 AGAAGATCTAGAACAGAACAAGCAAGTCAGGTCAATTCTCCACTCACTCCAGTGTGGTGGTAGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 AGAAGATCTAGAACAGAACAAGCAAGTCAGGTCAATTCTCCACTCACTCCAGTGTGGTGGTAGT 1550

Qy   661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTAGTGCATGCATGGCTTTCAGA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1551 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTAGTGCATGCATGGCTTTCAGA 1610

Qy   721 AGATCCATGGGCAAACATCTGTAGATGGACAGAGACCCGCTGGGTTCTTTTACAAGACAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 AGATCCATGGGCAAACATCTGTAGATGGACAGAGACCCGCTGGGTTCTTTTACAAGACAT 1670

Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACACAGTGCCTTTTTAGTGCATGCCTTTCAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 CCTTCTCAAATGGCAACGTCTTACTGAAGAACACAGTGCCTTTTTAGTGCATGCCTTTCAGA 1730

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1731 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1790

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGAATCTAGAAAAGAAAAGCAATCCAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1791 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGAATCTAGAAAAGAAAAGCAATCCAT 1850

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1851 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1910

Qy   1021 CCAGAAGACGGAAGCATGGCGATAACTTTGGGATAATTTAGTCCAAAA 1080
              |||||||||||||||||||||||||||||||||||||||||| |||||
Db   1911 CCAGAAGACGGAAGCATGGCGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1970

```
y    1081 ACTTGAAAAGAGTACAGCACAG--------------------------------------------------
          ||||||||||||||||||||||
o    1971 ACTTGAAAAGAGTACAGCAGCAGAAAACTGAAATAGCAGTTCAAGCTAAACAACCGGATGT 2030

y    1103 -------------------------------------------------------------
o    2031 GGAAGAGATTTGTCTAAAGGGCAGCATTTGTACAAGGAAAAACCAGCCACTCAGCCAGT 2090

y    1103 -------------------------------------------------------------
o    2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGGAAGGCGGTAAACCGTTTACTTCAAGA 2150

y    1103 -------------------------------------------------------------
o    2151 GCTGAGGGGCAAAGCAGCAGCCTAGCTCCTGACTCGACCACTATTGGAGCCTCTCCTAC 2210

y    1103 -------------------------------------------------------------
o    2211 TCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACT 2270

y    1103 -------------------------------------------------------------
o    2271 AGAAATGCCATCTCCTTGATGTTGGAGGTACCTGCTTCTCTGGCAGATTCAACCGGCTG 2330

y    1103 -------------------------------------------------------------
o    2331 GACAGAACTTACCGACTGGCTTTCTCTGCTTGATCAAGTTATATAAAATCACAGAGGGTGAT 2390

y    1103 -------------------------------------------------------------
o    2391 GGTGGGTGACTTGAGGATATCAACAGATGATCATCAAGCAGAGGCAACAATGCAGGA 2450

y    1103 -------------------------------------------------------------
o    2451 TTTGAACAGAGAGGCGTCCCAGTTGGAAGAACTCATTACCGCTGCCCAAAAATTTGAAAAA 2510

y    1103 -------------------------------------------------------------
o    2511 CAAGACCAGCAATCAGCAATCAGAGGCTAGAACAATCATTACGGATCGAATTGAAAGAATTCAGAA 2570

y    1103 -------------------------------------------------------------
o    2571 TCAGTGGGATGAAGTACAAGAACACACCTTCAGAACCGGAGGCAACAGTTGAATGAAATGTT 2630

y    1103 -------------------------------------------------------------
o    2631 AAAGGATTCAACACAATGGCTGAAGAGGCTAAGGAAGAAGCTGAGCAGTCTTAGGACAGGC 2690

y    1103 -------------------------------------------------------------
o    2691 CAGAGCCAAGCTTGAGTCATGGAAGGAGAGGGTCCCTATACAGTAGATGCAATCAAAAGAA 2750

y    1103 -------------------------------------------------------------
o    2751 AATCACAGAAAACCAAGCAGTTGGCCAAGACCTTCCGCCAGTGCAGACAAATGTGATGT 2810

y    1103 -------------------------------------------------------------
o    2811 GGCAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTGCAGATGATACCAGAGAAAGT 2870

y    1103 -------------------------------------------------------------
o    2871 CCACATGATAACAGAGAATATCAATGCCTCTTGGAAGCATTCATAAAAGGGTGAGTGA 2930

y    1103 ------------ACTCATAGATTACTGCAACAGTCCCCTGGA 1140
                         ||||||||||||||||||||||||||||
o    2931 GCGCAGAAGGCTGCTTTGCCCTTGGAGAAACATCATAGATTACTGCAACAGTCCCCTGGA 2990

y    1141 AAAGTTTCTTGCCCTGGCTTACAGAGAGTTCCTACAGGATGCTAC 1200
          |||||||||||||||||||||||||||||||||||||||||||||
o    2991 AAAGTTTCTTGCCCTGGCTTACAGAGAGTTCCTACAGGATGCTAC 3050

y    1201 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGAGGAGTAAAGAGCTGATGAAACAATGCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o    3051 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGAGGAGTAAAGAGCTGATGAAACAATGCA 3110

Qy   1261 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 3170

Qy   1321 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTTCCTGTTACAAAGACGTTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3171 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTT 3230

Qy   1381 GGATAACATGAACTTCAAGTGAGTGGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 GGATAACATGAACTTCAAGTGAGTGGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3290

Qy   1441 TTTGAAGCCAGTTCTGACCAGTGAAGGCTCGCACCTTTCTCGAGGAACTTCTGGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 TTTGAAGCCAGTTCTGACCAGTGAAGGCTCGCACCTTTCTCGAGGAACTTCTGGT 3350

Qy   1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCC 3410

Qy   1561 AGCAGTTCAGAAAGCAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3411 AGCAGTTCAGAAAGCAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 3470

Qy   1621 ACCTGTAATCATGAGTACTTCTTGAGACTGTACAGAATATTTCGACGAGCAGCCTTTGGA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3471 ACCTGTAATCATGAGTACTTCTTGAGACTGTACAGAATATTTCGACGAGCAGCCTTTGGA 3530

Qy   1681 AGGACTAGAGAAACTCTACCAGGAGCTGCCTCCTGAGGAGGTGCCCAGAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3531 AGGACTAGAGAAACTCTACCAGGAGCTGCCTCCTGAGGAGGTGCCCAGAA 3590

Qy   1741 TGTCACTCGGCTTCTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3591 TGTCACTCGGCTTCTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3650

Qy   1801 CCTGCACTCCGCTGACTCGGCAGAGAAAATAGATGAGACCCTGAAAGACTCCAGGAACT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3651 CCTGCACTCGCGCTGACTCGGCAGAGAAAATAGATGAGACCCTGAAAGACTCCAGGAACT 3710

Qy   1861 TCAAGAGGCCACGATGAGCTGACCTTCAAGCTGAGGTGACTGAGGTGATCAAGGGATC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3711 TCAAGAGGCCACGATGAGCTGACCTTCAAGCTGAGGTGACTGAGGTGATCAAGGGATC 3770

Qy   1921 CTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3771 CTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 3830

Qy   1981 GGCACTTCGAGGAGAATTGCGCCTCTGAAAGAGAAGCGTGAGCCAGTCAATGACCTTGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3831 GGCACTTCGAGGAGAATTGCGCCTCTGAAAGAGAAGCGTGAGCCAGTCAATGACCTTGC 3890

Qy   2041 TCGGCAGCTTACCACTTTGGGCATTCAGCTCTCAGCTCTCACCGTATAACCTCTGGAAGA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3891 TCGGCAGCTTACCACTTTGGGCATTCAGCTCTCAGCTCTCACCGTATAACCTCTGGAAGA 3950

Qy   2101 C 2101

Db   3951 C 3951
```

RESULT 12
AAD37260
ID   AAD37260 standard; DNA; 4414 BP.
XX
AC   AAD37260;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-3447.

XW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XW   Becker muscular dystrophy; ds.

XS   Homo sapiens.
XS   Unidentified.
XS   Chimeric.

XN   WO200183695-A2.

XD   08-NOV-2001.

XF   27-APR-2001; 2001WO-US013677.

XR   28-APR-2000; 2000US-0200777P.

XA   (XIAO/) XIAO X.

XI   Xiao X;

XR   WPI; 2002-049342/06.

XT   New dystrophin minigene for treating Duchenne or Becker muscular
XT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
XT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
XT   gene.

XS   Example 1; Page 65-66; 71pp; English.

XC   The present invention relates to an isolated nucleotide sequence encoding
XC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
XC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
XC   invention also relates to a recombinant adeno-associated virus (AAV)
XC   comprising dystrophin minigene operably linked to an expression control
XC   element. The dystrophin minigene in operable linkage with an expression
XC   control element, in a recombinant adeno-associated virus or retrovirus is
XC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XC   vector plasmid construct containing human dystrophin minigenes, a muscle
XC   creatine kinase (MCK) promoter and a small polyA signal sequence.

XQ   Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 U; 0 Other;

Query Match   53.0%; Score 1114; DB 6; Length 4414;
Best Local Similarity   97.0%;
Matches 1135; Conservative   0; Mismatches   35; Indels   0; Gaps   0;

Qy   1 GAGCTATGCCTACACAGAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCCTACACAGAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1716

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1776

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 1836

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy   241 CCAGTTTCATACTTCATACTCATGAGGGGTACAGCAGCCCATCAGGGCCGGGTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTTCATACTTCATCATGAGGGGTACAGCAGCCCATCAGGGCCGGGTGG 1956

Qy   301 TAATATTCTACAATTGGGAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

Qy   421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2136

Qy   481 GAAAGAGTTGATGACTGGCTAACAAAACAGAGAAAAACAGAGGAAAATGGGAGGAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTTGATGACTGGCTAACAAAACAGAGAAAAACAGAGGAAAATGGGAGGAAGA 2196

Qy   541 GCCTCTTGGACCTGATCTTGAAGACTTGAAGACGCCAAGTACAACAACATAAGGTGCTTCA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCTCTTGGACCTGATCTTGAAGACTTGAAGACGCCAAGTACAACAACATAAGGTGCTTCA 2256

Qy   601 AGAAGATCTAGAACAGAACAAGTCAGGGTCAATTCTCCACTCCACATGGTGGTGGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 AGAAGATCTAGAACAGAACAAGTCAGGGTCAATTCTCTCCACTCCACATGGTGGTGGTAGT 2316

Qy   661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 2376

Qy   721 AGATCCATGGGCAAACATCTGTAGATGGACAGAGACCCGTGGGTTCTTTTACAAGACAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTGGGTTCTTTTACAAGACAT 2436

Qy   781 CCTTCTCAAATGGCAACGTCTACTGACGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2437 CCTTCTCAAATGGCAACGTCTTACTGACGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 2496

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2497 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 2556

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCTGGATCGAGAAAAGAAAGAATCCAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2557 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCTGGATCGAGAAAAGAAAGCAATCCAT 2616

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2617 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2676

Qy   1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCAAAA 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2677 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 2736

Qy   1081 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA 1140
          |||||||||||||||||||||||||||||||||||||||| | |
Db   2737 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCAAGAGGCCACGGA 2796

Qy   1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGA 1170
          |||| ||
Db   2797 TGAGCTGGACCTCCAAGCTGCGCCAAGCTGA 2826

RESULT 13
ABK81997
ID   ABK81997 standard; DNA; 5417 BP.
XX
AC   ABK81997;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaR4-R23.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX

N WO20229056-A2.
X
D 11-APR-2002.
X
F 04-OCT-2001; 2001WO-US031126.
X
R 06-OCT-2000; 2000US-0238848P.
X
A (UNMI) UNIV MICHIGAN.
X
I Chamberlain JS, Harper SQ;
X
R WPI; 2002-435334/46.
X
T A composition for preparing therapeutic drugs, has a mini-dystrophin
T peptide comprising a specific number of spectrin-like repeat domains, or
T a nucleic acid sequence encoding the mini-dystrophin peptide.
X
S Disclosure; Fig 12; 145pp; English.
X
C The invention describes a composition comprising a mini-dystrophin
C peptide comprising a spectrin-like repeat domain, where the domain
C comprises n spectrin-like repeats, and contains no more than n spectrin-
C like repeats, where n is an even number between 4-24, or a nucleic acid
C encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
C polynucleotide encoding it is useful as a medicament, for preparing a
C drug for therapeutic application and in the preparation of a composition
C for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
C (DMD). This sequence represents a mini-dystrophin sequence of the
C invention
X
Q Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 U; 0 Other;

Query Match   52.5%; Score 1103.4; DB 6; Length 5417;
Best Local Similarity 99.9%; Pred. No. 5.8e-299;
Matches 1104; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGGAGAG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATACTCATGAGGGTACATGAGGGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1339 CCAGTTCATACTCATGAGGGTACATGAGGGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCCTAAATTGAGATCCTAAATTGGGAATCCTCAGGGTAGC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1459 AACTGAAGTACAAGAGCAGATGAATCCCTAAATTCAAGATGGGAATCCTCAGGGTAGC 1518

Qy   421 TAGCAGTGGAAAAACAAAGCAATTTACACAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1519 TAGCAGTGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAATGAGAGGAAGA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1579 GAAAGAGTTGAATGACTGGCTAACAACAAAAACAGAGAAAGAACAAGGAAAATGAGAGGAAGA 1638

Qy   541 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGGCCAAGTACAACAACATAAGGTGCTTCA 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGGCCAAGTACAACAACATAAGGTGCTTCA 1698

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTACTCCACTCCACATGGTGGTGGTAGT 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTACTCCACATGGTGGTGGTAGT 1758

Qy   661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTAAGGTATTGGG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTAAGGTATTGGG 1818

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTGGGTCTTTACAAGACAT 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTGGGTCTTTTACAAGACAT 1878

Qy   781 CCTTCTCCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1879 CCTTCTCCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAGATCAAATGAAAATGTT 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAATGAAAATGTT 1998

Qy   901 ATCAAGTCTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1999 ATCAAGTCTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy 1021 CCAGAGACCGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2119 CCAGAGACCGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 2178

Qy 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
         |||||||||||||||||||||||||
Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 14
ABK82005
ID ABK82005 standard; DNA; 11241 BP.
XX
AC ABK82005;
XX
DT 13-AUG-2002 (first entry)
XX
DE cDNA encoding human dystrophin, full length HDMD.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; human; ds.
XX
OS Homo sapiens.
XX
PN WO20229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031126.
XX
PR 06-OCT-2000; 2000US-0238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
DR WPI; 2002-435334/46.
XX
PT A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains, or
PT a nucleic acid sequence encoding the mini-dystrophin peptide.

```
XX
PS   Example 2; Fig 23; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a human dystrophin polynucleotide
CC   sequence used in the creation of the mini-dystrophin peptides of the
CC   invention
XX
SQ   Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 U; 0 Other;

Query Match    52.5%;   Score 1103.4;   DB 6;   Length 11241;
Best Local Similarity   99.9%;   Pred. No. 8.7e-299;
Matches 1104;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;
```

```
Qy     1 GAGCTATGCTCACACACAGGCGTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
         ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db  1099 GAGCTATGCTCACACACAGGCGCTATGTGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy    61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1218

Qy   121 TGAAGTAAACCTGACCGTTATCAAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 TGAAGTAAACCTGACCGTTATCAAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATACTTCATGAGGGGTACATGATTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTCATACTTCATGAGGGGTACATGATTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1398

Qy   301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1399 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459 AACTGAAGTACAAGAGCAGATGAATCCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGAGTTTAATGATCTCCAGAATCAGAAACT  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGAGTTTAATGATCTCCAGAATCAGAAACT 1578

Qy   481 GAAAGAGTTGAATGACTGCTAACAAAAACAGAGAAGAAGAAACAGGAAAATGGAGGAAGA  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579 GAAAGAGTTGAATGACTGCTAACAAAAACAGAGAAGAAGAAACAGGAAAATGGAGGAAGA 1638

Qy   541 GCCTCTTGACCTGATCTTGAGAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639 GCCTCTTGACCTGATCTTGAGAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 1698

Qy   601 AGAGAATCTAGAACAAGAACAAGTCAGGTCAATTCTCACTCACATGGTGGTGGTAGT  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699 AGAGAATCTAGAACAAGAACAAGTCAGGTCAATTCTCACTCACATGGTGGTGGTAGT 1758

Qy   661 TGATGAAATCTAGTGAGATGGAGATCACGGCCAACTGCTTGGAAGAACAACTTAAGGTATTGGG  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1759 TGATGAAATCTAGTGAGATGGAGATCACGGCCAACTGCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGTTCTTTTACAAGACAT  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGTTCTTTTACAAGACAT 1878
```

```
Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTGTGCATGGCTTTCAGA  840
         ||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db  1879 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTAGTGCATGGCTTTCAGA 1938

Qy   841 AAAAGAAGATGCAGTGAACAAGATTCACAACTGCCACAACTGGCTTTAAAGATCAAAATGAAATGTT  900
         ||||||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||
Db  1939 AAAAGAAGATGCAGTGAACAAGATTCACAACTGCCACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998

Qy   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAAGAAAAGAAGCAATCCAT  960
         |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGGATCTAGAAAAAGAAAAGAAGCAATCCAT 2058

Qy   961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2059 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy  1021 CCAGAAGACGGAAGCATGGCTGGATAAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119 CCAGAAGACGGAAGCATGGCTGGATAAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178

Qy  1081 ACTTGAAAAAGAGTACAGCACACAGACT 1105
         ||||||||||||||||||||||||||||
Db  2179 ACTTGAAAAAGAGTACAGCACACAGATT 2203
```

```
RESULT 15
ABK82002
ID   ABK82002 standard; DNA; 11443 BP.
XX
AC   ABK82002;
XX
DT   13-AUG-2002  (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaR9-R16.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200229056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Chamberlain JS,  Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drug, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 17; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 U; 0 Other;
```

```
Query Match      52.5%;  Score 1103.4;  DB 6;  Length 11443;
Best Local Similarity  99.9%;  Pred. No. 8.8e-299;
Matches 1104;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

y     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCACCGGAGCCCATT     60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCACCGGAGCCCATT   1158

y    61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   1218

y   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   1278

y   181 TGCTGAGGACACATTGCAAGCAGAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA   240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1279 TGCTGAGGACACATTGCAAGCAGAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA   1338

y   241 CCAGTTCATACTCATGAGGGGTACAGCCCATCAGGGCCCGGGTTGG   300
        ||||||||||||||||||||||||||||||||||||||||||||||
b  1339 CCAGTTCATACTCATGAGGGGTACAGCCCATCAGGGCCCGGGTTGG   1398

y   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAGATGGAAGA   360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAGATGGAAGA   1458

y   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   1518

y   421 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1519 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   1578

y   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGAGAGGAAGA   540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGAGAGGAAGA   1638

y   541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACACAACATAAGGTGCTTCA   600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1639 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTACACAACATAAGGTGCTTCA   1698

y   601 AGAAGATCTAGAACAAGAACAAGTCAGGGCAATTCTCTCACTACTCCACATGGTGTGGTAGT   660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1699 AGAAGATCTAGAACAAGAACAAGTCAGGGCAATTCTCTCACTACTCCACATGGTGTGGTAGT   1758

y   661 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTGGAAGACCGCTGGGTTCTTTACAAGACAT   720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGACCGCTGGGTTCTTTACAAGGTATTGGG   1818

y   721 AGATCGATGGGCAAAATCATCTGTAGATGACAGAGAGACCGCTGGGTTCTTTTACAAGACAT   780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1819 AGATCGATGGGCAAACATCATCTGTAGATGACAGAGAGACCGCTGGGTTCTTTTACAAGACAT   1878

y   781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTTAGTGCATGGCTTTCAGA   840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTTAGTGCATGGCTTTCAGA   1938

y   841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT   900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT   1998

y   901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCATCAGAAAAAGGAAATCCAT   960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGATCTAGAAAAAGGAAATCCAT   2058

y   961 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC   1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2059 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC   2118
```

```
Qy  1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA   1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119 CCAGAAGACCGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA   2178

Qy  1081 ACTTGAAAAGAGTACAGCACAGACT   1105
         |||||||||||||||||||||||||
Db  2179 ACTTGAAAAGAGTACAGCACAGATT   2203

Search completed: April  4, 2004, 18:19:26
Job time : 579.141 secs
```

M nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.
GenCore version 5.1.6

run on: April 4, 2004, 12:04:07 ; Search time 5477.46 Seconds
(without alignments)
16625.161 Million cell updates/sec

itle: US-09-845-416-2_COPY_900_3000
erfect score: 2101
equence: 1 gagctatgcctacacacagg........acctcagcactctggaagac 2101

coring table: IDENTITY.NUC
Gapop 10.0 , Gapext 1.0

earched: 3470272 seqs, 21671516995 residues

otal number of hits satisfying chosen parameters: 6940544

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: gb_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1293	61.5	8689	6	AX538622	AX538622 Sequence
2	1131	53.8	5952	6	AR304538	AR304538 Sequence
3	1131	53.8	5952	6	AX114289	AX114289 Sequence
4	1103.4	52.5	5417	6	AX538619	AX538619 Sequence
5	1103.4	52.5	11443	6	AX538624	AX538624 Sequence
6	1103.4	52.5	12057	6	AX538627	AX538627 Sequence
7	1103.4	52.5	13957	6	AX409637	AX409637 Sequence
8	1103.4	52.5	13957	6	AX538581	AX538581 Sequence
9	1103.4	52.5	13957	9	HUMDYS	M18533 Homo sapien
10	1103.4	52.5	13977	9	AR220819	AR220819 Sequence
11	1100.2	52.4	12446	9	HSDMDR	X14298 Human mRNA
12	997	47.5	5339	6	AX538620	AX538620 Sequence
13	996	47.4	5462	6	AX538621	AX538621 Sequence
14	929	44.2	13887	4	AF070485	AF070485 Canis fam
15	911.6	43.4	4402	6	E30220	E30220 Shortened d
16	866.6	41.2	3275	10	MUSDYS	M18025 Mouse dystr
17	866.6	41.2	13815	6	AX306153	AX306153 Sequence
18	866.6	41.2	13815	6	AX538582	AX538582 Sequence
19	866.6	41.2	13815	10	MUSDYSA	M68859 Mouse dystr
20	866.6	41.2	19307	6	AR093392	AR093392 Sequence
21	866.6	41.2	19307	6	AR142592	AR142592 Sequence
22	718.8	34.2	4075	6	E30221	E30221 Shortened d
23	714.6	34.0	4402	6	E30219	E30219 Shortened d
24	635.8	30.3	13575	9	GGDYS	X13369 Chicken mRN
25	624.2	29.7	630	9	HSDMDF1	X06179 Human fetal
26	527.8	25.1	1966	9	HSDMDA1	X06178 Human adult
27	514.6	24.5	1481	9	BC009242	BC009242 Homo sapi
28	409.2	19.5	3747	6	E30218	E30218 Shortened d
29	387	18.4	387	6	AX538612	AX538612 Sequence
30	361.4	17.2	10302	6	AX538583	AX538583 Sequence
31	361.4	17.2	10302	9	HSMUPS	X69086 H.sapiens m
32	350.4	16.7	10705	10	RNA029567	AJ002967 Rattus no
33	348	16.6	348	6	AX538611	AX538611 Sequence
34	344.6	16.4	1898	10	BC024140	BC024140 Mus muscu
35	344.4	16.4	3521	10	AF339031	AF339031 Danio rer
36	343	16.3	2341	10	BC062163	BC062163 Mus muscu
37	343	16.3	2394	10	BC050183	BC050183 Mus muscu
38	343	16.3	6045	6	A63605	A63605 Sequence 7
39	343	16.3	6045	6	AR281528	AR281528 Sequence 9
40	343	16.3	10320	6	A63607	A63607 Sequence 9
41	343	16.3	10320	6	AR281529	AR281529 Sequence
42	336.6	16.0	6059	6	AX107972	AX107972 Sequence
43	336.6	16.0	11096	6	AX538584	AX538584 Sequence
44	336.6	16.0	11096	10	MMY12229	Y12229 M.musculus
45	333	15.8	333	6	AX538589	AX538589 Sequence

ALIGNMENTS

RESULT 1
AX538622
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

Pred. No. is the number of results predicted by chance to have a

```
FEATURES          Location/Qualifiers
     source       1..8689
                  /organism="synthetic construct"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:32630"
                  /note="Synthetic"
ORIGIN

Query Match       61.5%;    Score 1293;  DB 6;  Length 8689;
Best Local Similarity  72.5%;  Pred. No. 3.8e-296;
Matches 2101;  Conservative  0;  Mismatches  0;  Indels 798;  Gaps  1;

Qy        1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCTACACGGAGCCCATT   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCTACACGGAGCCCATT 1158

Qy       61 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAGTCATTTGGCAGTTCATTTGATGAGAG  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1159 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAGTCATTTGGCAGTTCATTTGATGAGAG 1218

Qy      121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTTCTTTC 1278

Qy      181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA  240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA 1338

Qy      241 CCAGTTTCATACTCATGAGGGGTACATGATGAAGCTGATTGGAACAGGAAAATTATCAAGA  300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1339 CCAGTTTCATACTCATGAGGGGTACATGATGAAGCTGATTGGAACAGGAAAATTATCAAGA 1398

Qy      301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy      361 AACTGAAGTACAAGACGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1459 AACTGAAGTACAAGACGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy      421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1519 TAGCATGAAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy      481 GAAAGAGTGAATGACTGGCTAACAACAAAAAACAAGGAAAATGGAGAAGA  540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1579 GAAAGAGTTGAATGACTGGCTAACAACAAGAAGAACAAGGAAAATGGAGAAGA 1638

Qy      541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCAAGTACACACAACATAAGGTGCTTCA  600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCAAGTACACACAACATAAGGTGCTTCA 1698

Qy      601 AGAAGATCTAGAACAAGCAAGTCAGGTCAATTCTCACTCACAAGGTGGTGGTAGT  660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1699 AGAAGATCTAGAACAAGCAAGTCAGGTCAATTCTCACTCACAAGGTGGTGGTAGT 1758

Qy      661 TGATGAATCCAGTGAGATCACCGCAACTGCTGCTTTGAAGAACAACTTAAGGTATTGGG  720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1759 TGATGAATCCAGTGAGATCACCGCAACTGCTGCTTTGAAGAACAACTTAAGGTATTGGG 1818

Qy      721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT  780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy      781 CCTTTCAAATGGCAACGTCTTACTGAAGACAAGTGCCTTTTAGTGCATGGCTTTCAGA  840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1879 CCTTTCAAATGGCAACGTCTTACTGAAGACAAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy      841 AAAAGAGATGCAGTGAACAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT  900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1939 AAAAGAGATGCAGTGAACAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998

Qy      901 ATCAAGTCTTCAAAAACTGCCCGTTTAAAAGCGGATCTAGAAAGAGAAAAGAAAAGCAATCCAT  960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAGAGAAAAGAAAAGCAATCCAT 2058

Qy          GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAGTCAGTGAC 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAGTCAGTGAC 2118

Qy     1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAA 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTGTTGGGATAATTTAGTCCAAA 2178

Qy     1081 ACTTGAAAAGAGTACAGCACAG-                                        1102
            ||||||||||||||||||
Db     2179 ACTTGAAAAGAGTACAGCACAGATTTCACAGCAGCCTGACCTGAGCTTCCTGACTGACCAC 2238

Qy     1103 ------------------------------------                          1102
            
Db     2239 TATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGA 2298

Qy     1103 ------------------------------------                          1102
            
Db     2299 AACTGCCATCTCCAAACTAGAAATGCCATCTCCTTGATGTGGAGTACCTGCTCTGGC 2358

Qy     1103 ------------------------------------                          1102
            
Db     2359 AGATTTCAACCGGGCTTGACAGAACTTACCGACTGGCTTTCTGCTTGATCAAGTTAT 2418

Qy     1103 ------------------------------------                          1102
            
Db     2419 AAAATCACAGAGGGTGATGGTGGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA 2478

Qy     1103 ------------------------------------                          1102
            
Db     2479 GAAGGCAACAATGCAGAGATTTGAACAGAGAGGGTCCCCAGTTGGAAGAACTCATTACCGC 2538

Qy     1103 ------------------------------------                          1102
            
Db     2539 TGCCCAAAATTGAAAAACAAGAACAGCAATCAAGAGGCTAGAACAATCATTACGGATCG 2598

Qy     1103 ------------------------------------                          1102
            
Db     2599 AATTGAAAGAATTCAGATACAGTGGGATGGATGAAGTACAAGAACACCTTCAGAACCGGAGGCA 2658

Qy     1103 ------------------------------------                          1102
            
Db     2659 ACAGTTGAATGAAATGTTAAAGGTCAACACAATGGCTGGAAGCTAAGGAAGAAGCTGA 2718

Qy     1103 ------------------------------------                          1102
            
Db     2719 GCAGGTCTTAGGACAGGCCAGAGCCAAGCTTGAGTCATGGAGGAGGGTCCCTATACAGT 2778

Qy     1103 ------------------------------------                          1102
            
Db     2779 AGATGCAATCCAAAAGAAAATCACAGAACCAGCAGTGGCCAAGAACCTCCGCCAGTG 2838

Qy     1103 ------------------------------------                          1102
            
Db     2839 GCAGACAAATGTAGATGTGGCAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTGC 2898

Qy     1103 ------------------------------------                          1102
            
Db     2899 AGATGATACCAGAAAAGTCCACATGATAACAGAGAATATCAATGCCTCTTGGAGAAGCAT 2958

Qy     1103 ------------------------ACTCATAGATTACTGCAACA                  1122
                                    ||||||||||||||||
Db     2959 TCAATAAAGGGTGAGTGAGCGAGAGGCTGCTTTGGAGAAAACTCATAGATTACTGCAACA 3018

Qy     1123 GTTCCCCCTGGACCTGGACCTTCTTGGCTTTACAGAAGCTGAAACAACTGCCAA 1182
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3019 GTTCCCCCTGGACCTGGACCTTCTTGGCTTTACAGAAGCTGAAACAACTGCCAA 3078

Qy     1183 TGTCCTACAGGATGCTACCCGTAAGGAAGAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGA 1242
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
REFERENCE   1 (bases 1 to 5952)
AUTHORS     Xiao,X. and Liu,P.X.
TITLE       Method and vector for producing and transferring trans-spliced
            peptides
JOURNAL     Patent: US 6544786-A 1 08-APR-2003;
FEATURES             Location/Qualifiers
source               1..5952
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match          53.8%;  Score 1131;  DB 6;  Length 5952;
Best Local Similarity 68.6%;  Pred. No. 1.2e-257;
Matches 2101;  Conservative 0;  Mismatches 0;  Indels 960;  Gaps 1;

Qy      1 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    891 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 950

Qy     61 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    951 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 1010

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGTATTATCGTTGCGTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1011 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGTATTATCGTTGCGTTCTTTC 1070

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1130

Qy    241 CCAGTTCATACTCATGAGGGTACATGATTGGATTTGACAGCCCATCAGGGCCGGTGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1131 CCAGTTCATACTCATGAGGGTACATGATTGGATTTGACAGCCCATCAGGGCCGGTGGTTGG 1190

Qy    301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1191 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA 1250

Qy    361 AACTGAAGTACAAGAGCGAGATGAATCTCCTAAATTCAAGATGGGAATGGGCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1251 AACTGAAGTACAAGAGCGAGATGAATCTCCTAAATTCAAGATGGGAATGGGCTCAGGGTAGC 1310

Qy    421 TAGCATGGAAAAACAAATTTACATGAGTTTAATGAACAGAGAAAATGGAGGAAGAAAACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 TAGCATGGAAAAACAAATTTACATGAGTTTAATGAACAGAGAAAATGGAGGAAGAAAACT 1370

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1371 GAAAGAGTTGAATGACTGGCTAACAAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1430

Qy    541 GCCTCTTGGACCTGATCTTGAAGACCTAAAGCGCAATTCTCCTCACATGTGGTGGTAGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 GCCTCTTGGACCTGATCTTGAAGACCTAAAGCGCAATTCTCCTCACATGTGGTGGTAGT 1490

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATCTCTCACCTCTCCACACAACTTAAGGTATTGGG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATCTCTCACCTCTCCACACAACTTAAGGTATTGGG 1550

Qy    661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGCTTTGGAAGAACACCGCTGGGGTTCTTTTACAAGACAT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 AGATCGATGGGCAAACATCTGTAGATGGAGATCACGCAACTGCTGCTTGCTTTGGAAGAACACCGCTGGGGTTCTTTTACAAGACAT 1670

Qy    721 AGATCGATGGGCAAACATCTGTAGATGGAGATCACGCAACTGCTGCTTGCTTTGGAAGAACACCGCTGGGGTTCTTTTACAAGACAT 780

Qy    781 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1730

Qy    841 AAAAGAAGATGCAGTGAACAGAGATTCACACAACTGCTTAAAGATCAAAATGAAAATGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   3079 TGTCCTACAGGATGCTACCCGTAAGGAAAAGCTCCTAGAGGAACTCCAAGGGAGTAAAAGA 3138

Qy   1243 GCTGATGAAAACAATGGCAAGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCA 1302
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3139 GCTGATGAAAACAATGGCAAGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCA 3198

Qy   1303 CAACCTGATGAAAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCCGATGATGCAGT 1362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3199 CAACCTGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 3258

Qy   1363 CCTGGTTACAAAGAGACGTTTGATAACATGAACTTCAAGTGGAGTGAACTTCGGGAAAAAGTC 1422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3259 CCTGGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGGAAAAAGTC 3318

Qy   1423 TCTCAACATTAGTAGTCCCATTTGGAAGCCAGTTCTGACCACCAGGCGTCTGGAGAGCGTCTGACCACCTTTC 1482
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3319 TCTCAACATTAGTAGTCCCATTTGGAAGCCAGTTCTGACCACCAGGCGTCTGACCAGGCGTCTGACCACCTTTC 3378

Qy   1483 TCTGCAGGAACTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCAGGCACC 1542
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3379 TCTGCAGGAACTTCTCGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGCAGGCACC 3438

Qy   1543 TATTGGAGGCGACTTTCAGAGCAGTGTACATAGGGCCTTCAAGAG 1602
          |||||||||||||||||||||||||||||||||||||||||||||
Db   3439 TATTGGAGGCGACTTTCAGAGCAGAACGAGAACGACATAGGGCCTTCAAGAG 3498

Qy   1603 GGAATTGAAAAACTAAAGAACCTGTAATCAGTGTCTCTGAGACTGTACGAATATTTCT 1662
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3499 GGAATTGAAAAACTAAAGAACCTGTAATCAGTGTCTCTGAGACTGTACGAATATTTCT 3558

Qy   1663 GACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCGCTGCCTCC 1722
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3559 TGAGGAGCAGCCTTTGGAAGGACTCCAGAGAAACTCTACCAGGAGCCGCTGCCTCC 3618

Qy   1723 TGAGGAGGAGCCAGAATGTCACTCGGCTTCTACGAGAAAGCAGGCTGAGGAGGTCAATAC 1782
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3619 TGAGGAGGAGCCAGAATGTCACTCGGCTTCTACGAGAAAGCAGGCTGAGGAGGTCAATAC 3678

Qy   1783 TGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCT 1842
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3679 TGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCT 3738

Qy   1843 TGAAAGACTCCAGGAGACTTCAAGAGAGGGATCCTGCAGCCGGTGGCCACGGATGAGCTGCGCCAAGC 1902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3739 TGAAAGACTCCAGGAGACTTCAAGAGAGGGATCCTGCAGCCCGTGGGCCACGGATGAGCTGCGCCAAGC 3798

Qy   1903 TGAGGTGATCAAGGGGAATCCTGGGCGATCTCGGGCGAAATTGCGCCTCTGAAAGAGAACGTGAG 1962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3799 TGAGGTGATCAAGGGGAATCCTGGGCGATCTCGGGCGAAATTGCGCCTCTGAAAGAGAACGTGAG 3858

Qy   1963 TCACCTTCGAGAAAGTCAAGGCACTTCGAGGGCCTCTGAAAGAGAACGTGAG 2022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3859 TCACCTTCGAGAAAGTCAAGGCACTTCGAGGCCTCTGAAAGAGAACGTGAG 3918

Qy   2023 CCAGCTCAATGACCTTGCTCGCCGGCCAGTCTGACCACTTTGGGCATTCAGCTCTCACCGTATAA 2082
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3919 CCACGTCAATGACCTTGCTCGCCGGCCAGTCTGACCACTTTGGGCATTCAGCTCTCACCGTATAA 3978

Qy   2083 CCTCAGCACTCTGGAAGAC 2101
          |||||||||||||||||||
Db   3979 CCTCAGCACTCTGGAAGAC 3997

RESULT 2
AR304538
LOCUS       AR304538            5952 bp    DNA         linear   PAT 12-JUN-2003
DEFINITION  Sequence 1 from patent US 6544786.
ACCESSION   AR304538
VERSION     AR304538.1  GI:31693691
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
```

```
Qy  1731 AAAGAAAGATGCAGTGAACAAGAGATTCACCACAACTGGCTTTTAAAGATCAAATGAAATGTT 1790

Db        AAAGAAAGATGCAGTGAACAAGAGATTCACCACAACTGGCTTTTAAAGATCAAATGAAATGTT 1102

Qy   901 ATCAAGTCTTCAAAAAACTGGCCGTTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT  960
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  1791 ATCAAGTCTTCAAAAAACTGGCCGTTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1850

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAGTCAGTGAC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1851 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAGTCAGTGAC 1910

Qy  1021 CCAGAGACGGAAGCATGCTGGATAACTTGCCGGTGTTGGGATAATTAGTTCCAAAA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1911 CCAGAGACGGAAGCATGCTGGATAACTTGCCGGTGTTGGGATAATTAGTTCCAAAA 1970

Qy  1081 ACTTGAAAGAGTACAGCACAG----------- 1102
         |||||||||||||||||||||
Db  1971 ACTTGAAAGAGTACAGCACAGGAAAACTGAAAACTAGCAGTCAAACATGAAAACCGGATGT 2030

Qy  2031 GGAAGAGATTTGTCTAAAGGGCAGCATTGTACAAGGAAAAACCAGCCACTCAGCCAGT 2090

Db        GGAAGAGATTTGTCTAAAGGGCAGCATTGTACAAGGAAAAACCAGCCACTCAGCCAGT 1102

Qy  2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGGAAGGCGGTAAACCGTTTACTTCAAGA 2150

Db        GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGGAAGGCGGTAAACCGTTTACTTCAAGA 1102

Qy  2151 GCTGAGGGCAAAGCAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTAC 2210

Db        GCTGAGGGCAAAGCAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTAC 1102

Qy  2211 TCAGACTGTTACTCTGGTGACACAACCTGGTTACTAAGGAAAACTGCCATCTCCAAACT 2270

Db        TCAGACTGTTACTCTGGTGACACAACCTGGTTACTAAGGAAAACTGCCATCTCCAAACT 1102

Qy  2271 AGAAAATGCCATCTCCTGATGTGTTGGAGTACCTGCTCTGGCAGATTCAACCGGGCTTG 2330

Db        AGAAAATGCCATCTCCTGATGTGTTGGAGTACCTGCTCTGGCAGATTCAACCGGGCTTG 1102

Qy  2331 GACAGAACTTACCGGACCTGGCTTTCTCTGCTTCAGTTATAAATCAGACAGAGGGTGAT 2390

Db        GACAGAACTTACCGGACCTGGCTTTCTCTGCTTCAGTTATAAATCAGACAGAGGGTGAT 1102

Qy  2391 GGTGGGTGGACCTTGAGGATATCAACGAGATGATGATCAAGCGAAGGCAACAATGCGAGGA 2450

Db        GGTGGGTGGACCTTGAGGATATCAACGAGATGATGATCAAGCGAAGGCAACAATGCGAGGA 1102

Qy  2451 TTTGGAACAGAGGGCGTCCCAGTTGGAAGAACTCATTACCGCTGCCCAAAATTTGAAAAA 2510

Db        TTTGGAACAGAGGGCGTCCCAGTTGGAAGAACTCATTACCGCTGCCCAAAATTTGAAAAA 1102

Qy  2511 CAAGACCACGCAATCAAGAGGCTAGAACAATCATTACCGGATCGAAAGAATTCAGAA 2570

Db        CAAGACCACGCAATCAAGAGGCTAGAACAATCATTACCGGATCGAAAGAATTCAGAA 1102

Qy  2571 TCAGTGGGATGAAGTGAAGTACAAGAACACCTTCAGAACCTTCAGAGTTGAATGAAAATGTT 2630

Db        TCAGTGGGATGAAGTGAAGTACAAGAACACCTTCAGAACCTTCAGAGTTGAATGAAAATGTT 1102

Qy  2631 AAAGGATTCAACACAACATGCTGGAAGCTAAGGAAGAGAGCTGGAGTCTTAGGACAGGC 2690

Db        AAAGGATTCAACACAACATGCTGGAAGCTAAGGAAGAGAGCTGGAGTCTTAGGACAGGC 1102

Qy  2691 CAGAGCCAAGCTTCAGTCATGGAGTCATGATGCAATCCAAAAGAA 2750

Db        CAGAGCCAAGCTTCAGTCATGGAGTCATGATGCAATCCAAAAGAA 1102

Qy  2751 AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTTGGCAGACAAATGTAGATGT 2810

Db        AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTTGGCAGACAAATGTAGATGT 1102

Qy  2811 GGCAAATGACTTGGCCCCTGAAACTTCTCCGGGATTATTCTGCAGATGATACACCAGAAAAGT 2870

Db        GGCAAATGACTTGGCCCCTGAAACTTCTCCGGGATTATTCTGCAGATGATACACCAGAAAAGT 1102
```

```
Qy     2101 C 2101
Db     3951 C 3951

RESULT 3
AX114289
LOCUS       AX114289              5952 bp    DNA        linear     PAT 11-MAY-2001
DEFINITION  Sequence 1 from Patent WO0129243.
ACCESSION   AX114289
VERSION     AX114289.1  GI:14031259
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Paul,X.L. and Xiao,X.
TITLE       Method and vector for producing and transferring trans-spliced
            peptides
JOURNAL     Patent: WO 0129243-A 1 26-APR-2001;
            DALHOUSIE UNIVERSITY (CA); UNIV. OF PITTSBURGH OF THE COMMONWEALTH
            SYSTEM OF HIGHER EDUCATION (US)
FEATURES             Location/Qualifiers
     source          1..5952
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    2897..2898
                     /note="S4 junction site"
     misc_feature    3198..3199
                     /note="S2 junction site"
ORIGIN

Query Match             53.8%;  Score 1131;  DB 6;  Length 5952;
Best Local Similarity   68.6%;  Pred. No. 1.2e-257;
Matches 2101;  Conservative    0;  Mismatches    0;  Indels  960;  Gaps    1;
```

```
Qy      1 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCACCTCAC  50
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    891 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCACCTCAC  940

Qy     51 CCATTGGCAGTTCATTGAGAG  60
            |||||||||||||||||||||
Db    941 CCATTGGCAGTTCATTGAGAG  950
```

```
Qy   1103 -------------------------------------------------------------------- 1102
Qb   2511 CAAGACCAGCAATCAAGAGGCTAGACAATCATTACGACGTTGAATTGAAAGAATTCGAA 2570

Qy   1103 -------------------------------------------------------------------- 1102
Qb   2571 TCAGTGGGATGAAGTACAAGAACACACCTTCAGAACCGGAGCAACAGTTGAATGAAAATGTT 2630

Qy   1103 -------------------------------------------------------------------- 1102
Qb   2631 AAAGGATTCAACACAATGGCTAGGAAGCTAAGGAAGAAGCTGAGCAGGTCTTAGGACAGGC 2690

Qy   1103 -------------------------------------------------------------------- 1102
Qb   2691 CAGAGCCAAGCTTGAGTCATGGAAGGAGGAGGTCCCTATACAGTAGATGCAATCCAAAAGAA 2750

Qy   1103 -------------------------------------------------------------------- 1102
Qb   2751 AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCCCAGTGGCAGACAAATGTAGATGT 2810

Qy   1103 -------------------------------------------------------------------- 1102
Qb   2811 GGCAAATGACTTGGCCCCTGGAAACTTCTCCGGATTATTCTGCAGATGATACCAGAAAAGT 2870

Qy   1103 -------------------------------------------------------------------- 1102
Qb   2871 CCCACATGATAACAGAGAGAATATCAATGCCTCTTGGAGAAGCATTCATAAAAGGGTGAGTGA 2930

Qy   1103 -------------ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1140
Qb   2931 GCGAGAGGCGCTGCTTTGGAAGAAACTCATAGATTACTGCAACAGTTCCCCCTGGACCTCGA 2990
               |||||||||||||||||||||||||||||||||||||||||

Qy   1141 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1200
Qb   2991 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 3050
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1201 CCGTAAGGAAAGGCTCCTAGAAGACTTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 1260
Qb   3051 CCGTAAGGAAAGGCTCCTAGAAGACTTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 3110
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1261 AGACCCCAAGTGAAATTGAAGCTGAAGCTCACACAGATGTTATCACAACCTGATGAAAACAG 1320
Qb   3111 AGACCCTCCAAGTGAAATTGAAGCTGAAGCTCACACAGATGTTATCACAACCTGATGAAAACAG 3170
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1321 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCCTGTTTACAAAGACGTTT 1380
Qb   3171 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCCTGTTTACAAAGACGTTT 3230
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1381 GGATAACATGAAGACTTCAAGTGAGTGAACTTCGGAGAAAAAGTCTCTCAACATTAGGTCCCA 1440
Qb   3231 GGATAACATGAAGACTTCAAGTGAGTGAACTTCGGAGAAAAAGTCTCTCAACATTAGGTCCCA 3290
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1441 TTTGGAAGCCAGTTCCACCAGTCCAGTGGAAGCCGTCTGCACCTTTCTCTGCAGGAACTTCTGT 1500
Qb   3291 TTTGGAAGCCAGTTCCACCAGTCCAGTGGAAGCCGTCTGCACCTTTCTCTGCAGGAACTTCTGT 3350
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1501 GTGGCTACAGTCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTTCC 1560
Qb   3351 GTGGCTACAGTCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTTTCC 3410
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1561 AGCAGTTCAGAGAGCAGAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1620
Qb   3411 AGCAGTTCAGAGAGCAGAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 3470
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1621 ACCTGTAAACATGAGTACTCTTGACACGTACGAATATTTCTGACAGAGCCTTTGGA 1680
Qb   3471 ACCTGTAAACATGAGTACTCTTGACACGTACGAATATTTCTGACAGAGCGCCTTTGA 3530
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1681 AGGACTAGAGAAACTTCACCAGGAGAGCTCCTCGAGGAGAGAGAGAGTGGTGAAAGA 1740
Qb   3531 AGGACTAGAGAAACTTCACCAGGAGAGCTCCTCCTGGAGGAGAGAGAGAGTGGTGAAAGA 3590
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1741 TGTCACTCGGCTTCTACGAAAGCAGCAGGTCGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
```

```
Db   3591 TGTCACTCGGCTTCTACGAAAGCAGCAGGTCGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 3650
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1801 CCTGCACTCCGCTGACTGGCAGGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1860
Db   3651 CCTGCACTCCGCTGACTGGCAGGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 3710
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1861 TCAAGAGGGCACGGATGAGCTGGACCTCAAGCTGCGCGGAGCTGAGGTGATCAAGGGATC 1920
Db   3711 TCAAGAGGGCACGGATGAGCTGGACCTCAAGCTGCGCGGAGCTGAGGTGATCAAGGGATC 3770
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1921 CTGGCAGCCCGTGGGCGATCTCTTCCCAAGATCACCTCGAGAAGAAGTCAA 1980
Db   3771 CTGGCAGCCCGTGGGCGATCTCTTCCCAAGATCACCTCGAGAAGAAGTCAA 3830
               ||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1981 GGCACTTCGAGGAGAAATTGCGCCCTGAAAGAGAACGTGAGCCGTCAATGACCTTGC 2040
Db   3831 GGCACTTCGAGGAGAAATTGCGCCCTGAAAGAGAACGTGAGCCGTCAATGACCTTGC 3890
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2041 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCTCGGAAGA 2100
Db   3891 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCTCGGAAGA 3950
               |||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2101 C 2101

Db   3951 C 3951
```

```
RESULT 4
AX538619
LOCUS       AX538619                5417 bp    DNA         linear   PAT 23-NOV-2002
DEFINITION  Sequence 39 from Patent WO0229056.
ACCESSION   AX538619
VERSION     AX538619.1  GI:25271163
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 39 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..5417
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

Query Match             52.5%;  Score 1103.4;  DB 6;  Length 5417;
Best Local Similarity   99.9%;  Pred. No. 4.1e-251;
Matches 1104;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy      1 GAGCTATGCTACACAGGCTGCTTATGTCACCACCTGACCCTACACGGAGCCCATT 60
Db   1099 GAGCTATGCTACACAGGCTGCTTATGTCACCACCCCTCTGACCCTACACGGAGCCCATT 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATTGGAGAG 120
Db   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATTGGAGAG 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    121 TGAAGTAAACTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db   1219 TGAAGTAAACTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    241 CCAGTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGAGCGCGGGTTG 300
```

FEATURES　　　　THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
source　　　　Location/Qualifiers
　　　　　　　1..11443
　　　　　　　/organism="synthetic construct"
　　　　　　　/mol_type="unassigned DNA"
　　　　　　　/db_xref="taxon:32630"
　　　　　　　/note="Synthetic"

ORIGIN

Query Match　　　52.5%;　　Score 1103.4;　　DB 6;　　Length 11443;
Best Local Similarity　99.9%;　　Pred. No. 4.1e-251;
Matches 1104;　Conservative　0;　Mismatches　0;　Indels　1;　Gaps　0;

```
Qy        1 GAGCTATGCTACACACAGCGTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT       60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1099 GAGCTATGCTACACACAGCGTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT     1158

Qy       61 TCCTTCACAGCATTGGAAGCTCCTAGAAGCAAGTCATTCGCAGTTCATTCATGGAGAG      120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1159 TCCTTCACAGCATTGGAAGCTCCTGAAGCAAGTCATTCGCAGTTCATTCATGGAGAG     1218

Qy      121 TGAAGTAAACCTGGACCCTTATCAAACGCTTTAGAAGAAGTATTATCGTCGTGCTTCTTC      180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTCGTGCTTCTTC     1278

Qy      181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGTAAGTGGTGAAAGA      240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGTAAGTGGTGAAAGA     1338

Qy      241 CCAGTTCATACTCATGAGGGGTACAGCCCATCAGGGCCGGGTTGG      300
            ||||||||||||||||||||||||||||||||||||||||||||
Db     1339 CCAGTTCATACTCATGAGGGGTACAGCCCATCAGGGCCGGGTTGG     1398

Qy      301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA      360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA     1458

Qy      361 AACTGAAGTACAAGAGCAGATGAATCTTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC      420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1459 AACTGAAGTACAAGAGCAGATGAATCTTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC     1518

Qy      421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT      480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1519 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT     1578

Qy      481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAGAACAACTAAGGAAGA      540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAGAACAACTAAGGAAGA     1638

Qy      541 GCCTCTTGGACCTGATCTTGAAGACTCAAGACCTCAAGGGTCAATTCTCTCACTCAAGGTGCTTCA      600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1639 GCCTCTTGGACCTGATCTTGAAGACTCAAGACCTCAAGGGTCAATTCTCTCACTCAAGGTGCTTCA     1698

Qy      601 AGAAGATCTAGAACAAGAACAAGACAAGAACAGCAAGAACAAGACCAAGTCAGGGTAGT      660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1699 AGAAGATCTAGAACAAGAACAAGACAAGAACAGCAAGAACAAGACCAAGTCAGGGTAGT     1758

Qy      661 TGATGAATCTAGTTGGAGATTACAAGAACAACTTAAGGTATTGGG      720
            |||||||||||||||||||||||||||||||||||||||||||
Db     1759 TGATGAATCTAGTTGGAGATTACAAGAACAACTTAAGGTATTGGG     1818

Qy      721 AGATCCGATGGGCAAACATCTGTAGATGGGCAAACATCTGTAGATGGACAAGAGACCGCTGGGTTCTTTTACAAGACAT      780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1819 AGATCCGATGGGCAAACATCTGTAGATGGACAAGAGACCGCTGGGTTCTTTTACAAGACAT     1878

Qy      781 CCTTCTCAAATGGCAAACGTCTTACTGACAGTGCCTTTTTAGTGCATGGCTTCAGA      840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1879 CCTTCTCAAATGGCAACGTCTTACTGACAGTGCCTTTTTAGTGCATGGCTTTCAGA     1938

Qy      841 AAAAGAAGATGCAGTGAACAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT      900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1939 AAAAGAAGATGCAGTGAACAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT     1998
```

```
o      1339 CCAGTTCATACTCATGAGGGGTACAGCCCATCAGGGCCGGGTTGG     1398

/       301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA      360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o      1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA     1458

/       361 AACTGAAGTACAAGAGCAGATGAATCTTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC      420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o      1459 AACTGAAGTACAAGAGCAGATGAATCTTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC     1518

/       421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT      480
o      1519 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT     1578

/       481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAGAACAACTAAGGAAGA      540
o      1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAGAACAACTAAGGAAGA     1638

/       541 GCCTCTTGGACCTGATCTTGAAGACTCAAGACCTCAAGGGTCAATTCTCTCACTCAAGGTGCTTCA      600
o      1639 GCCTCTTGGACCTGATCTTGAAGACTCAAGACCTCAAGGGTCAATTCTCTCACTCAAGGTGCTTCA     1698

/       601 AGAAGATCTAGAACAAGAACAAGACAAGAACAGCAAGAACAAGACCAAGTCAGGGTAGT      660
o      1699 AGAAGATCTAGAACAAGAACAAGACAAGAACAGCAAGAACAAGACCAAGTCAGGGTAGT     1758

/       661 TGATGAATCTAGTTGGAGATTACAAGAACAACTTAAGGTATTGGG      720
o      1759 TGATGAATCTAGTTGGAGATTACAAGAACAACTTAAGGTATTGGG     1818

/       721 AGATGCGGATGGGCAAACATCTGTAGATGACAGAGACCGCTGGGTTTAAAGACCGTTGGCTTCA      780
o      1819 AGATGCGGATGGGCAAACATCTGTAGATGACAGAGACCGCTGGGTTTAAAGACCGTTGGCTTCA     1878

/       781 ATCAAGTCTTCAAAAACTGTATTCACTCAAACAAGATCTTCTTTTGCCGGTGTTGGGATAATTTAGTCCAAAA      840
o      1879 ATCAAGTCTTCAAAAACTGTATTCACTCAAACAAGATCTTCTTTTGCCGGTGTTGGGATAATTTAGTCCAAAA     1938

/       841 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTTGCCGGTGTTGGGATAATTTAGTCCAAAA     1020
o      1999 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTTGCCGGTGTTGGGATAATTTAGTCCAAAA     2058
```

```
o      1081 ACTTGAAAAGAGTACAGCACAGACT 1105
                  ||||||||||||||||||  |
o      2179 ACTTGAAAAGAGTACAGCACAGATT 2203
```

RESULT 5

LOCUS　　　AX538624　　　　11443 bp　　DNA　　　linear　　PAT 23-NOV-2002
DEFINITION　Sequence 44 from Patent WO0229056.
ACCESSION　AX538624
VERSION　　AX538624.1　　GI:25271175
KEYWORDS　.
SOURCE　　synthetic construct
ORGANISM　synthetic construct
　　　　　　artificial sequences.
REFERENCE　1
AUTHORS　　Chamberlain,J.S. and Harper,S.Q.
TITLE　　　Mini-dystrophin nucleic acid and peptide sequences
JOURNAL　　Patent: WO 0229056-A 44 11-APR-2002;

RESULT 6

```
RESULT 6
AX538627
LOCUS       AX538627              12057 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 47 from Patent WO0229056.
ACCESSION   AX538627
VERSION     AX538627.1  GI:25271181
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 47 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..12057
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

Query Match             52.5%;  Score 1103.4;  DB 6;  Length 12057;
Best Local Similarity   99.9%;  Pred. No. 4.1e-251;
Matches 1104;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    0;
```

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTCAGTTCATTGCAGTTCATTGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTCAGTTCATTGCAGTTCATTGGAGAG 1218

Qy    121 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy    181 TGTGAGACACATTGCAAGCACAAGGAGATTTCTAATGTGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGTGAGACACATTGCAAGCACAAGGAGATTTCTAATGTGTGGAAGTGGTGAAAGA 1338

Qy    241 CCAGTTCATACTCCATGAGGGGTACAGATGAATTGACAGCCATCAGGGCCGGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTCATACTCCATGAGGGGTACAGATGAATTGACAGCCATCAGGGCCGGGTTGG 1398

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy    361 AACTGAAGTAGTACAGAGAGATGAATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTAGTACAGAGAGATGAATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy    421 TAGCATGGAAAAACAAAGCAATTTACATAGAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAGAAGAACAAGAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAGAAGAACAAGAAGA 1638

Qy    541 GCCTCTTGGACCTGATCTTGAAGACCTGAAGACCCAAGTACAACAACATAAGGTGCTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACCTGAAGACCCAAGTACAACAACATAAGGTGCTTCA 1698

Qy    601 AGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACTCCTCACATGGTGGTGGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACTCCTCACATGGTGGTGGTAGT 1758

Qy    661 TGATGAATCTAGTGGAGTAGTGGGCAAACATCTGTAGTGGACAGAAGAACCGCTGGGTTCTTTTACAAGACAT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAGTAGTGGGCAAACATCTGTAGTGGACAGAAGAACCGCTGGGTTCTTTTACAAGACAT 1818

Qy    721 AGATCGATGGGCAAACATCTGTAGTGGACAGAAGAACCGCTGGGTTCTTTTACAAGACAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGTGGACAGAAGAACCGCTGGGTTCTTTTACAAGACAT 1878

Qy    781 CCTTCTCAAATGGCAACGTCTTACTGAAGACACCAGTGCCTTTTAGTGCATGGCATTTCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTTACTGAAGACACCAGTGCCTTTTAGTGCATGGCATTTCAGA 1938

Qy    841 AAAAGAAGATGCAGTAACAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTAACAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998

Qy    901 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAGAAAAAGCAATCCAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAGAAAAAGCAATCCAT 2058

Qy    961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy   1021 CCAGAAGACGGAGCATGGCTGCTGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAAGACGGAGCATGGCTGCTGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 2178

Qy   1081 ACTTGAAAAGAGTACAGCACAGACT 1105
          ||||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGATT 2203
```

RESULT 7

```
RESULT 7
AX409637
LOCUS       AX409637              13957 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 2284 from Patent WO0229103.
ACCESSION   AX409637
VERSION     AX409637.1  GI:21442342
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2284 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Homo sapiens"
ORIGIN

Query Match             52.5%;  Score 1103.4;  DB 6;  Length 13957;
```

```
Db    2119  CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178

Qy    1081  ACTTGAAAAGAGTACAGCACAGACT 1105
            |||||||||||||||||||||||
Db    2179  ACTTGAAAAGAGTACAGCACAGATT 2203
```

RESULT 8
AX538581

```
LOCUS          AX538581    13957 bp    DNA    linear    PAT 23-NOV-2002
DEFINITION     Sequence 1 from Patent WO0229056.
ACCESSION      AX538581
VERSION        AX538581.1  GI:25271086
KEYWORDS       .
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
  AUTHORS      Chamberlain,J.S. and Harper,S.Q.
  TITLE        Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL      Patent: WO 0229056-A 1 11-APR-2002;
               THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Query Match          52.5%;  Score 1103.4;  DB 6;  Length 13957;
Best Local Similarity 99.9%;  Pred. No. 4.1e-251;
Matches 1104;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy      1  GAGCTATGCCTACACACAGCGTCGTTATGTCACCACCTCTGACCCACCGGAGCCCATT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099  GAGCTATGCCTACACACAGCGTCGTTATGTCACCACCTCTGACCCACCGGAGCCCATT 1158

Qy     61  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGAGGAG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGAGGAG 1218

Qy    121  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 1278

Qy    181  TGCTGAGGACACATTGCAAGCACCAAGGAGATTCTAATGATGTGGAAGTCGTGAAAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279  TGCTGAGGACACATTGCAAGCACCAAGGAGATTCTAATGATGTGGAAGTCGTGAAAGA 1338

Qy    241  CCAGTTCATACTCATGATGATGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339  CCAGTTCATACTCATGATGATGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy    301  TAATATTCTACAATGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399  TAATATTCTACAATGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy    361  AACTGAAGTACAAGAGCAGCAGAATGAATCTCCTAAATTCAAGAGTTTAATGGATCTCCAGAATCAGAAACT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459  AACTGAAGTACAAGAGCAGCAGAATGAATCTCCTAAATTCAAGAGTTTAATGGATCTCCAGAATCAGAAACT 1518

Qy    421  TAGCATGAGAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519  TAGCATGAGAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy    481  GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAGAACAGAAGAAGAACAAGGAAAATGGAGGAAGA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579  GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAGAACAGAAGAAGAACAAGGAAAATGGAGGAAGA 1638

Qy    541  GCCTCTTGGACCTGATCTTGAAGACTACAACGCCAAGTACAACACAACATAAGGTGCTTCA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639  GCCTCTTGGACCTGATCTTGAAGACTACAACGCCAAGTACAACACAACATAAGGTGCTTCA 1698

Qy    601  AGAGACATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGTGTGTGAGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699  AGAGACATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGTGTGTGAGT 1758

Qy    661  TGATGAATCTAGTGGAGATCACGCAGATCAACAACTGGCTTGCTTGGAAGAACAACTTAAGGTATTGGG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759  TGATGAATCTAGTGGAGATCACGCAGATCAACAACTGGCTTGCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy    721  AGATCGATGGGCAAACATCTGTAGAATGGACAGAAGACCGCTGGTTCTTTTTACAAGACAT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819  AGATCGATGGGCAAACATCTGTAGAATGGACAGAAGACCGCTGGTTCTTTTTACAAGACAT 1878

Qy    781  CCTTCTCAAATGGCAAGCTCTTACTGAAGAACAGTGCCCTTTTAGTGCATGCTTTCAGA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879  CCTTCTCAAATGGCAAGCTCTTACTGAAGAACAGTGCCCTTTTAGTGCATGCTTTCAGA 1938

Qy    841  AAAAGAAGATGCAGTGAACAAGATTCACCAACTGGCTTAAAGATCAAAATGAAAATGTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939  AAAAGAAGATGCAGTGAACAAGATTCACCAACTGGCTTAAAGATCAAAATGAAAATGTT 1998

Qy    901  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058

Qy    961  GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGGTAACAAGAATAAGTCAGTGAC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059  GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGGTAACAAGAATAAGTCAGTGAC 2118

Qy   1021  CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639  GCCTCTTGGACCTGATCTTGAAGACTACAACGCCAAGTACAACACAACATAAGGTGCTTCA 1698
```

```
Qy    601 AGAAGATCTAGACAAGAACAAGTCAGGTCAATTCTCCACTCACATGGTGGTCGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGACAAGAACAAGTCAGGTCAATTCTCCACTCACATGGTGGTCGTAGT 1758

Qy    661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy    721 AGATCCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGAGACAT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGAGACAT 1878

Qy    781 CCTTCTCAAATGGCAACGTCTTACTACAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTTACTACAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy    841 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998

Qy    901 ATCAAGTCTCAAAAACTGGCCCGTTTAAAAGCGGGATCTAGAAAAGGGATAAAAAGCAATCCAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTCAAAAACTGGCCCGTTTAAAAGCGGGATCTAGAAAAGGGATAAAAAGCAATCCAT 2058

Qy    961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy   1021 CCAGAAGACGGAAGCATGGCGAAGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAAGACGGAAGCATGGCGAAGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 2178

Qy   1081 ACTTGAAAAGAGTACAGCACAGACT 1105
          |||||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGATT 2203

RESULT 9
HUMDYS
LOCUS       HUMDYS           13957 bp  mRNA    linear   PRI 25-MAY-2000
DEFINITION  Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION   M18533 M17154 M18026 M20250
VERSION     M18533.1  GI:181856
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1699)
  AUTHORS   Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and
            Kunkel,L.M.
  TITLE     Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
            preliminary genomic organization of the DMD gene in normal and
            affected individuals
  JOURNAL   Cell 50 (3), 509-517 (1987)
  MEDLINE   87273512
  PUBMED    3607877
REFERENCE   2  (bases 1678 to 3830)
  AUTHORS   Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
  TITLE     Conservation of the Duchenne muscular dystrophy gene in mice and
            humans
  JOURNAL   Science 238 (4825), 347-350 (1987)
  MEDLINE   88018015
  PUBMED    3659917
REFERENCE   3  (bases 1 to 13957)
  AUTHORS   Koenig,M., Monaco,A.P. and Kunkel,L.M.
  TITLE     The complete sequence of dystrophin predicts a rod-shaped
            cytoskeletal protein
  JOURNAL   Cell 53 (2), 219-226 (1988)
  MEDLINE   88194521
  PUBMED    3282674
COMMENT     On May 25, 2000 this sequence version replaced gi:340693.
```

Draft entry and computer-readable sequence kindly provided by
M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
determined by the size of the deleted DNA segment. Deletions found
in different patients were from positions 302-2200, 473-1168,
1691-1810, and 1169-3011.

```
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xp21.3-p21.1"
                     /tissue_type="muscle"
                     /dev_stage="fetus"
                     /note="G00-119-850"
     gene            1..13957
                     /gene="DMD"
     CDS             209..11256
                     /gene="DMD"
                     /codon_start=1
                     /product="dystrophin"
                     /protein_id="AAA53189.1"
                     /db_xref="GI:181857"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSLMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVKEEHFQLHHQMHYSQ
                     QITVSLAQGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFPSQHLEAPEDKSFGSSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTLQAGEISNDVEVVKDQFHTHEGYMDLTAH
                     QGRVGNILQLGSKLIGTGKLSEDEETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLKELNDWLTKTEERTRKMEEEPLGPDLEDLKRQVQOHKVLQEDLEEQVRVN
                     SLTHMVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWVLLQDILLKMQRLTE
                     EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLQKLAVLKADLEKKKQSMGKLIYSLK
                     QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTQPSLTQTTVM
                     ETVTVTTREQILVKHAQELPPPQQKKRQITVDSEIRKRLDVDITELHSWITRSEA
                     VLQSPEFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSAQALVEQMVNEGVNA
                     DSIKASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLEQMTTAENWLKIQ
                     PTTPSEPTAIKSQLKICKDEVNRLSGLQPIERLKIQSIALKEKGQGPMFLDADFVAF
                     TNHFKQVFSDVQAREKELQTIPDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
                     EIMEQRLGEIQALQSSLQEQQSGLYYLSTTVKEMSKKAPSEISRKYQSEFEEIEGRWK
                     KLSSOLVERCQKLEEQMNKLRKIQNHIQTLKKWMAEVDVKGSGEVKPALQDSELKKQL
                     KQCRLLVSDIQTIQPSLNSVNEGGQKINEAEPPFASRLETELKELNTQWDHMCQQVY
                     ARKEALKGGLEKTVSLQKDLSEMHEWMTQAEEEYLERDFEYKTPDELQKAVEEMKRAK
                     EEAQQKEAKVKLLTESVNSVIAQAPPVAQEALKKELETLTTNYQWLCTRLNGKCKTLE
                     EVWACWHELLSYLEKANKWLNEVEFKLKTTENIPGGAEEISEVLDSLENLMRHSEDNP
                     NQIRILAQTILTDGGVMDELINEELETFNSRWRELHEEAVRRQKLLEQSIQSAQETEKS
                     LHLIQESLTFIDKQLAAYIADKVDAAQMPEAQKIQSDLTSHEISLEMKKHQGKEA
                     AQRVLSQIDVAQKKLQDVSMKFRLFQKPANFELRLQSKMILDEVKOMHLPALETKSVE
                     QEVVQSQLNHCVNLYKSLSEVKSEVEMVIKTGRQIVQKKQTENPKELDERVTALKLHY
                     NELGAKVTERKQQLEKCLKLSRKVRKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
                     VAWGKATQKEIEKQKVHLKSITEVGEALKTVLGKKETLVEDKLSLLNSNWIAVTSRAE
                     EWLNLLEYQKHMETFDQNVDHITKWIIQADTLLDESEKKKPQQKEDVLKRLKAELND
                     IRPKVDSTRDQAANLMANRGDHCKKLVEPQISELNHRFAAISHRIKTGKASIPLKELE
                     QFNSDIQKLLEPLEAEIQQGVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQQRITDER
                     KREEIKIKQLLQTKNALKDLRSQRRKALEISHQWYQYKRQADDLLKCLDDIEKKL
                     ASLPEPRDERKIKEIDREIQKKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWRE
                     IESKPAQFRRLNFAQIHTVREETMVVMTEDMPLEISVYPSTYLTEITHVSQALLEVEQ
                     LLNAPDLCAKDFEDLFKQEESLKNIKDSLQQSSGRIDIIHSKTAALQSATPVERVKL
                     QEALSQLDFQWEKVNKMYKDRQGFPDRSVEKWRRFHYDIKIFNQWLITEAEQFLRKTQI
                     PENWEHAKYKWYLKELQDGIGQRQTVRTLNATGEEIIQQSSKTDASILQERLGSLNL
                     RWQEVCKQLSDRKKRLEEQKNILSEFQRDLNEFVLWLEEADNIASIPLEPGEKQQLKE
                     KLEQVKLLVEELPLRQGIIKQLNETGGPVLVSAPISPEEQKLKENLKQTNLQWIKVS
                     RALPEKQGEIEAQIKDLGQLKLEDLEEQLNHLLLWLSPIRNQLEIYNQPNQEGPFD
                     VQETEIAVQAKQPDVEEILSKGOHLYKEKPATOPVKRKLEDLSSEWKAVNRLLQELRA
                     KQPDLAPGLTTIGASPTQTVTLVTQPVVTKETAISKLEMPSSLMLEVPALADFNRAWT
                     ELTDWLSLLDQVIKSQRVMVGDLEDINEMIIKQKATMQDLEQRRPQLEELITAAQNLK
                     NKTSNQEARTIIDRIERIQNQWDESHLQRNRQQLNEMLKDSTQWLEEKEEAEQVL
                     GQARAKLESWKEGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSAD
                     DTRKVHMITENIINASWRSIHKRVSEREAALEETHRLIQQFPLDLEKFIAWLTEAETTA
                     NVLQDATRKERLLEDSKGVKELMKQWQDLQGIEAHTDVYHNLDENSQKILRSLEGSD
                     DAVLLQRRLDNMNFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELLVWLQLKDDEL
                     SRQAPIGGDFPAVCKQNDVHRAFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLYQ
```

EPRELPPERAQNVTRLLRKQAEEVNTEWEKLNLHSADWQRKIDETLERLQELQEATD
ELDKLRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEIAPLKENVSHVNDLARQL
TTLGIQLSPYNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVQGP
WERAISPNKVPYINHEIQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA
LCLDLLSLSAACCALDQINLKQNDQPMDILQINCLTTIYDRLEQEHNNLVNVPLCVD
MCLNWLLNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRL
GLLIHDSIQIPRQLGEVASFGGSNIEPSVRSCFQPANKPEIEAALFLDWMRLEPQSM
VWLPVLHRVAAAETAKHQAKCNICKECPIIGFRYRSKHFNYDICQSCFFSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPMGYLPVQTVLEGDNMET
PVTLINFWPVDSAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSVLNDSISPNESID
DEHLLIQHYCQSLNQDSPLSQPRSPAQILISLESERGELERILADLEEBNRNLQAEY
DRLKQQHEHKGLSPLPSPPEMMPTSFQSPRDAELIAEAKLRQHKGRLEARMQILEDH
NKQLESQLHRLRQLLEQPQAEAKVNGTTVSSPSTSLQRSDSSQPMLLRVVGSQTSDSM
GEEDLLSPPQDTSTGLEEVMEQLANSFPSSSRGRNTPGKPMREDTM"

RIGIN

Query Match   52.5%;   Score 1103.4;   DB 9;   Length 13957;
Best Local Similarity 99.9%;   Pred. No. 4.1e-251;
Matches 1104;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

y   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

y   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG 1218

y   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

y   181 TGCTGAGGACACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1279 TGCTGAGGACACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

y   241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1339 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 1398

y   301 TAATATTCTACAATTGGAAGCTGATTGGAACAGGAAAAATTATCAGAAGAAGATGAAGA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1399 TAATATTCTACAATTGGAAGCTGATTGGAACAGGAAAAATTATCAGAAGAAGATGAAGA 1458

y   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

y   421 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1519 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

y   481 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAGAACAAGGAAAATGAGAGAAGA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1579 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAGAACAAGGAAAATGAGAGAAGA 1638

y   541 GCCTCCTTGGACCTGATCTTGAAGACAAGCCAAGTACAACCAAGTACAACAACATAAGGTGCTTCA   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1639 GCCTCCTTGGACCTGATCTTGAAGACAAGCCAAGTACAACCAAGTACAACAACATAAGGTGCTTCA 1698

y   601 AGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTGTGTGGGTTGGTAGT   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1699 AGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTGTGTGGGTTGGTAGT 1758

y   661 TGATGAATCTAGTGTGAGAAAACAAGCAACTCACGCAACTCGCTGCTTGGAAGAACAACTTAAGGTATTTGGG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1759 TGATGAATCTAGTGTGAGAAAACAAGCAACTCACGCAACTCGCTGCTTGGAAGAACAACTTAAGGTATTTGGG 1818

y   721 AGATCGATGGGCAAACATCTGTAGATGACAGAAGAACCGCTGGAAGTAAGCTGGTTCTTTTTACAAGACAT   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1819 AGATCGATGGGCAAACATCTGTAGATGACAGAAGAACCGCTGGAAGTAAGCTGGTTCTTTTTACAAGACAT 1878

y   781 CCTTCCAAATGGCAACGTCTTACTGAAGACAACAGTGCCTTTTTAGTGCATGGCTTTCAGA

RESULT 10
AR220819
LOCUS   AR220819   13977 bp   DNA   linear   PAT   26-SEP-2002
DEFINITION   Sequence 60 from patent US 6426186.
ACCESSION   AR220819
VERSION   AR220819.1   GI:23327696
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 13977)
  AUTHORS   Jones,K.A., Volkmuth,W. and Walker,M.G.
  TITLE   Bone remodeling genes
  JOURNAL   Patent: US 6426186-A 60 30-JUL-2002;
FEATURES   Location/Qualifiers
     source   1..13977
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN

Query Match   52.5%;   Score 1103.4;   DB 6;   Length 13977;
Best Local Similarity 99.9%;   Pred. No. 4.1e-251;
Matches 1104;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   0;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

Qy   241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1339 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGAAGCTGATTGGAACAGGAAAAATTATCAGAAGAAGATGAAGA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1399 TAATATTCTACAATTGGAAGCTGATTGGAACAGGAAAAATTATCAGAAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1519 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAGAACAAGGAAAATGAGAGAAGA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAGAACAAGGAAAATGAGAGAAGA 1638

Qy   541 GCCTCCTTGGACCTGATCTTGAAGACAAGCCAAGTACAACCAAGTACAACAACATAAGGTGCTTCA   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1639 GCCTCCTTGGACCTGATCTTGAAGACAAGCCAAGTACAACCAAGTACAACAACATAAGGTGCTTCA 1698

Qy   601 AGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTGTGTGGGTTGGTAGT   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1699 AGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTGTGTGGGTTGGTAGT 1758

Qy   661 TGATGAATCTAGTGTGAGAAAACAAGCAACTCACGCAACTCGCTGCTTTGCTTGGAAGAACAACTTAAGGTATTTGGG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1759 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 1818

Qy   721 AGATCGATGGGCAAACATCTGTAGATGACAGAAGAACCGCTGGAAGTAAGCTGGTTCTTTTTACAAGACAT   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1819 AGATCGATGGGCAAACATCTGTAGATGACAGAAGAACCGCTGGAAGTAAGCTGGTTCTTTTTACAAGACAT 1878

Qy   781 CCTTCCAAATGGCAACGTCTTACTGAAGACAACAGTGCCTTTTTAGTGCATGGCTTTCAGA   840

Qy   841 AAAAGAAGATGCAGTGAACAAGAACAACAGTGCCTTTTTAAAGATCAAATCAAAATGAAATGTT   900

Qy   901 ATCAAGTCTTCAAAAACTGCGGTTTAAAAGCCGGATCAGAAAAGAAAAAGCAATCCAT   960

Qy   961 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC   1020

Qy 1021 CCAGAGACGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTAGTCCAAAA   1080

Qy 1081 ACTTGAAAAGAGTACAGCACAGACT   1105

Db 1879 CCTTCCAAATGGCAACGTCTTACTGAAGACAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1938

Db 1939 AAAAGAAGATGCAGTGAACAAGAACAACAGTGCCTTTTTAAAGATCAAATCAAAATGAAATGTT 1998

Db 1999 ATCAAGTCTTCAAAAACTGCGGTTTAAAAGCCGGATCAGAAAAGAAAAAGCAATCCAT 2058

Db 2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 2118

Db 2119 CCAGAGACGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTAGTCCAAAA 2178

Db 2179 ACTTGAAAAGAGTACAGCACAGATT 2203

```
Db  1459 AACTGAAGCTACACAGAGCAGCAGAATGCATGATGGGGATGCCTCAGGGTAGC 1518
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Zy   421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAAACT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1519 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATGATCTCCAGAATCAGAAAACT 1578

Zy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAGAAGAAATGGAGGAAGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAGAAGAAATGGAGGAAGA 1638

Zy   541 GCCTCTTGGACCTGATCTTGAAGACTTGAAGAAACGCCAAGTACAACATAAGGTGCTTCA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639 GCCTCTTGGACCCTGATCTTGAAGACTTGAAGAAACGCCAAGTACAACACATAAGGTGCTTCA 1698

Zy   601 AGAAGATCTAGAACAAGAACAGTCAGGGTCAATTCTCTCACTCACTCCACATGGTGGTAGT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699 AGAAGATCTAGAACAAGAACAGTCAGGGTCAATTCTCTCACTCACTCCACATGGTGGTAGT 1758

Zy   661 TGATGAATCTAGTGGAGTATCACGCAACTGCTGCTTGCTTGGAAGAACAACTTAAGGTATTGGG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1759 TGATGAATCTAGTGGAGTATCACGCAACTGCTGCTTGCTTGGAAGAACAACTTAAGGTATTGGG 1818

Zy   721 AGATCGATGGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819 AGATCGATGGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

Zy   781 CCTTCTCAAATGCCAAGCTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1879 CCTTCTCAAATGCCAAGCTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1938

Zy   841 AAAAGAAGATGCAGTGACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1939 AAAAGAAGATGCAGTGACAAGATTCACCACAACTGGCTTAAAGATCAAAATGAAATGTT 1998

Zy   901 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAGCAGGATCTAGAAAGAAAAGAAATCCAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1999 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCAGGATCTAGAAAGAAAAAAGCAATCCAT 2058

Zy   961 GGGCAAACTGTATTCACTCAAACAAGATCTCTTCTTCCCCGTGTTGGGATAATTAAGTCAGTGAC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2059 GGGCAAACTGTATTCACTCAAACAAGATCTCTTCTTCAACAGATCTGAAGAGAATAAGTCAGTGAC 2118

Zy  1021 CCAGAAGACGGAAGCGGAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTAGTCCAAAA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119 CCAGAAGACGGAAGCATGGCGGATGGCTGGATAACTTTGCCCGTGTTGGGATAATTAGTCCAAAA 2178

Zy  1081 ACTTGAAAAGAGTACAAGCACAGCAGACT 1105
         |||||||||||||||||||||||||||||
Db  2179 ACTTGAAAAGAGTACAAGCACAGCACAGATT 2203
```

RESULT 11

LOCUS   HSDMDR   12446 bp   mRNA   linear   PRI 12-SEP-1993
DEFINITION Human mRNA for dystrophin.
ACCESSION   X14298
VERSION   X14298.1 GI:30845
KEYWORDS   Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 12446)
AUTHORS   Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
            Davies,K.E.
TITLE   Two human cDNA molecules coding for the Duchenne muscular dystrophy
            (DMD) locus are highly homologous
JOURNAL   Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE   89345106
PUBMED   2668885
REFERENCE   2 (bases 1 to 12446)
AUTHORS   Rosenthal,A.

TITLE   Direct Submission
JOURNAL   Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
            der DDr, zentralinstitut fuer Molekularbiologie, Robert-Roessle
            Str.10, 1115 Berlin Buch, DDR
COMMENT   see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
            M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449,
            3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES   Location/Qualifiers
     source   1..12446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X chromosomal, Xp21."
                     /tissue_type="muscle"
                     /dev_stage="fetal and adult."
     CDS   99..11156
                     /note="unnamed protein product; dystrophin (AA 1 - 3685)"
                     /codon_start=1
                     /protein_id="CAA32479.1"
                     /db_xref="GI:30846"
                     /db_xref="GOA:P11532"
                     /db_xref="SWISS-PROT:P11532"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSILMVITSLFQVLPQQVSIEAIQEVEMLPRPPKKVTKEEHFQLHHQMHYSQ
                     QITVSLAQGYERTSSPKPFKSYAYTQAAYVTTSDPTRSPFPSQHLEAPEDKSFGSSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTLQAGGEISNDVEVVKDQFHTHEGYMMDLTAH
                     QGRVGNILQLGSKLIGTGKLSEDEETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLKELNDMLTKTEERTRKMEEEPLGPDLEDLKRQVQQHKVLQEDLEQEQVRVN
                     SLTHMVVVESSGDHATEALEEQLKVLGDRWANICRWTEDRRVLLQDILLKWQRLTE
                     EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLQKLAVILKADLEKKQSMGKLYSIK
                     QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTTQPSLTQTTVM
                     ETVTTVTTREQILVKHAQELPPPPPQKKRQITVDSEIRKRLDVDITELHSWITRSEA
                     VLQSPEFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSGQALVEQMVNEGVNA
                     DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQELEQMTTAENWLKIQ
                     PTTPSEPTAIKSQLKICKDEVNRLSGLGPQIERLKIQSIALKEKGQGMFLDADFVAF
                     TNHPKQVFSDVQAREKELQTIFDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
                     EIMEQRLGELQALQSSLQEQQSGLYYLSTTVKEMSKKAPSEISRKYQSBEFEIEGRWK
                     KLSSOLVEHCQKLEBQMNKLRKIQNHIQTLKKMWAEBVDVFLKEBWPALGDSILKKQL
                     KQCRLLVSDIQTIQPSLNSVNEGGQKIKKNBEAEPFASRLETELKELNTQWDRMCQQVY
                     ARKEALKGGLEKTVSLQKDLSEMEHWMTQAEEYLERDFEYKTPDELQKAFEMEKRAK
                     EEAQQKEAKVKLLITESVNSVIACAPPVAQEALKKELETLTTRYQWLCTRLNGCKCTLE
                     EVWACWHELLSYLEKANKWLNEVFFKLTTENIPGGAEBEISEVIDSLENLMRHSEDNP
                     NQIRLAQCTLTDGGVWMDELINEEELETFNSRWRELHEBAVRRQLLEQSIQSAQETENS
                     LHLIQESLTFIDKQLAAYIADKVDAAOMPQEAOKIQSDLTSHEISLEMKKHNQGKEA
                     AQRVLSQIDVAQKKLQDVSMKFRLFQKPANFEQRLQBSKMILDEVKMHLPALETKSVE
                     QEVVQSQLNHCVNLYKSLSEVKSVEMVIKTGRQIVQKKQTENPKELDERVTALKLHY
                     NELGAKVTERKQQLEKCLKLSRKMRKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
                     VAWGKATQKEIEKQKVHLKSITEVGEALKTIVGKKETIVEDKLSLLNSNWIAVTSRAE
                     EWLNLLLEYQKHMETFDQNVDHITKWIIQADTLLDESBKKKPQQKEDVLKRLKAELND
                     IRPKVDSTRDQAANLMANHGDHCRKLVEPQISELNHRFAAISHRIKTGKASIPLKELE
                     QFNSDIQKLLEPLEABIQQGVNLKEBDPKDMWEDNEGTVKELLQSKRWRE
                     KSEEIKIKQQLLQTKHNALKDLRSQRRKLALEISHQWYQYKRQADDLLKCLDDIEKKL
                     ASLPEPRDRKIKEIDRELQKKEELNAVRRQAEBLSDGAAMAVEPTQIQLSKRWRE
                     IESKFAQFRRLNFAQIHTVREETMMWMEDNEMPLEISVYPSTYLTEITHVSQALLEVG
                     LLNAPDLCAKDFEDLFKQEESLKNIKDSLQQSSGRIDIIHSKKTAALQSATPVERVKL
                     QEALSQLDFQWEKVNKMYKDRQGRFDRSVEKNRRFHYDIKIFNQWLTEAEQFLRKTQI
                     PENWEHAKYKWYLKELQDGIGQRQTVVRTLNATGEEIIQQSSKTDASLQEKLGSLNL
                     RWQEVCKQLSDRKKRLEEQKNILSEFQRDLNEFVLWLEEADNIASIPLEPGKEQQLKE
                     KLEQVKLLVEBLPLRQGILKQLNETGDGPVLVSAPISPEBEQDKLENKLKQTNLQWIKVS
                     RALPEKGGEIEAQIKDLGQLEKKLEDLNLLLKTPNKGQEAQKLEKERE
                     VKEETEIAVQAKQPDVEEILSKGQHLYEKPATQPVKRKLEDLSSEWKAVNRLLQELRA
                     KQPDLAPGLTTIGASPTQVTLVTQPVVTKETAISKLEMPSSIMLEVPALADFNRAWT
                     ELTDWLSLLDQVIKSQRVMVGDLEDINMIIKQKATMQDLEQRPQLELEIITAAQNLK
                     NKTSNQBARTIITDRIERIQNQWDEVQDEHLLQNRRQQLNEMLKDSTQWLEAKEEAEQVL
                     GQARAKLEBWKEGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSAD
                     DTRKVHMITENINASWRSIHKRVSEBRAALEBETHLQQFFPLDFLAWLTEAETTA
                     NVLQDATRKERLLBDSKGVKELMKQWQDLQGEIEAHTDVYHNLDENSQKILRSLEGSD
                     DAVLQRRLDNWFKWSELRKKLNIRSHLBASSDQWKRLHLSLQELLVWLQLKDDEL
                     SRQAPIGGDFPAVQKQNDVHRAFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLYQ
                     EPRELPPEERAQNVTRLLRKQAEBVNTEWEKLNLHSADWQRKIDETLERLQELQEATD

```

```
         ELDLKLRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEIAPLKENVSHVNDLARQL
         TTLGIQLSPYNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVGGP
         WERAISPNKVPYYINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA
         LCLDLLSLSAACDALDQHNLKQNDQPMDILQLINCLTTIYDRLEQEHNNLVNVPLCVD
         MCLNWLLNVYDTGTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRL
         GLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIBAALFLDWMRLEPQSM
         VWLPVLHRVAAAETAKHQAKCNICKECIIGFRYRSLKHFNYDICCSCFFSGRVAKGH
         KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLEGDNMET
         PVTLINFWPVDSAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSYLNDSISPNESID
         DEHLLIQHYCQSLNQDSPLSQPRSPAQLISLESEERGELERILADLEEENRNLQAEY
         DRLKQQHEHKGLSPLPSPPEMMFTSPQSPRDAELIABAKLLRQHKGRLEARMQILEDH
         NKQESQLHRLRQLLEQPQAEAKVNGTTVSSPSTSLQRSDSSQPMLLRVVGSQTSDSM
         GEEDLLSPQDTSTGLEEVMEQLNNSFPSSRGRNTPGKPMREDTM"
```
RIGIN

Query Match    52.4%;    Score 1100.2;    DB 9;    Length 12446;
Best Local Similarity   99.7%;    Pred. No. 2.4e-250;
Matches 1102;    Conservative   0;    Mismatches   0;    Indels   3;    Gaps   0;

```
Y        1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGACCCTCACGCGAGCCCATT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      989 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGACCCTCACGCGAGCCCATT 1048

Y       61 TCCTTCCACAGCAGTTGGAAGCTCCTGAAGACAAGTCATTGGAGTCATTGGAGGATGGAG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1049 TCCTTCCACAGCAGTTGGAAGCTCCTGAAGACAAGTCATTGGAGTCATTGGAGGATGGAG 1108

Y      121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1109 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 1168

Y      181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1169 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1228

Y      241 CCAGTTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCCGGGTGG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1229 CCAGTTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCCGGGTGG 1288

Y      301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1289 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1348

Y      361 AACTGAAGTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1349 AACTGAAGTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1408

Y      421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1409 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1468

Y      481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAATGAGAGGAAGA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1469 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAATGAGAGGAAGA 1528

Y      541 GCCTCTTGGACCTGATCTTGAAGACAGACCTAAGATGGTGGTGCTTCA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||
b     1529 GCCTCTTGGACCTGATCTTGAAGACAGACCTAAGATGGTGGTGCTTCA 1588

Y      601 AGAAGATCTAGAACAAGAGACAAGTCAGGGTCAATTCTCACTCACTTCTCCTGGAAGAACAACAAGTCAATTCTCC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1589 AGAAGATCTAGAACAAGAGACAAGTCAGGGTCAATTCTCACTCACTTCTCCTGGAAGAACAACAAGTCAATTCTCC 1648

Y      661 TCATGAATCAGTGTAATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1649 TCATGAATCAGTGTAATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1708

Y      721 AGATCGATGGGCAAACATCTGTAGAACCGCTGGGTTCTTTTACAAGACAT  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||
b     1709 AGATCGATGGGCAAACATCTGTAGAACCGCTGGGTTCTTTTACAAGACAT 1768

Y      781 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTAGTGCATGGCTTTCAGA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1769 CCTGCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1828

Qy     841 AAAAGAAGATGCAGTGAACAAGAGATTCACAACAACTGGCTTAAAGATCAAAATGAAATGTT  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1829 AAAAGAAGATGCAGTGAACAAGAGATTCACAACAACTGGCTTAAAGATCAAAATGAAATGTT 1888

Qy     901 ATCAAGTCTTCAAAAAACTGCCGTTTTAAAAAGCGGATCTGAGAAAGAAAAAAGCAATCCAT  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1889 ATCAAGTCTTCAAAAAACTGCCGTTTTAAAAAGCGGATCTGAGAAAGAAAAAAGCAATCCAT 1948

Qy     961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1949 GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2008

Qy    1021 CCAGAGACCGAAGCATGGCGATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2009 CCAGAGACCGAAGCATGGCGATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2068

Qy    1081 ACTTGAAAAGAGTACAGCACAGACT 1105
           |||||||||||||||||||||||||
Db    2069 ACTTGAAAAGAGTACAGCACAGATT 2093
```

RESULT 12
AX538620
```
LOCUS       AX538620                5339 bp    DNA        linear       PAT 23-NOV-2002
DEFINITION  Sequence 40 from Patent WO0229056.
ACCESSION   AX538620
VERSION     AX538620.1  GI:25271166
KEYWORDS    synthetic construct
SOURCE      synthetic construct
  ORGANISM  artificial sequences.
            1
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 40 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
  source          1..5339
                  /organism="synthetic construct"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:32630"
                  /note="Synthetic"
ORIGIN
```

Query Match   47.5%;   Score 997;   DB 6;   Length 5339;
Best Local Similarity 100.0%;   Pred. No. 7.8e-226;
Matches 997;   Conservative   0;   Mismatches   0;   Indels.   0;   Gaps   0;

```
Qy    1105 TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1164
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1546 TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1605

Qy    1165 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1224
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1606 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1665

Qy    1225 CTCCAAGGGAGTAAAAGAGTGATGAAACAATGCAAGACCTCCAAGGTGAAATTGAAGC 1284
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1666 CTCCAAGGGAGTAAAAGAGTGATGAAACAATGCAAGACCTCCAAGGTGAAATTGAAGC 1725

Qy    1285 TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCTGAGATCCCTGGA 1344
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1726 TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCTGAGATCCCTGGA 1785

Qy    1345 AGGTTCCGATGATGATGCAGTCCAATGTCACCCGTAAGATAACATGAACATGAAGTGGAG 1404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1786 AGGTTCCGATGATGATGCAGTCCAATGTCACCCGTAAGATAACATGAACATGAAGTGGAG 1845

Qy    1405 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1846 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1905
```

```
Qy  1465 GAAGCGTCGCACCTTCTCTGCGAGGAACTTCTGGTCTGGCTACAGCTGAAGAGATGATGA 1524
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1906 GAAGCGTCGCACCTTCTCTGCGAGGAACTTCTGGTGTGGCTACAGCTGAAGAGATGATGA 1965

Qy  1525 ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTCAGCAGTTCAGAGCAGAAGCGATGT  1584
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1966 ATTAAGCCGGCAGGCACCTATTGGAGGCGGACTTCAGCAGTTCAGAGCAGAAGCGATGT  2025

Qy  1585 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1644
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2026 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2085

Qy  1645 GACTGTACGAGAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGA 1704
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2086 GACTGTACGAGAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGA 2145

Qy  1705 GCCCAGAGAGCTGCCTCCTGAGGAGAGACTCCAGAAGTCACTCGGCTTCACGAAAGCA 1764
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2146 GCCCAGAGAGCTGCCTCCTGAGGAGAGACTCCAGAATGTCACTCGGCTTCACTACGAAAGCA 2205

Qy  1765 GGCTGAGGAGGTCAATACTGAGTGCGAAAAATTGAACCTGCACTGGGCGATCTCCT 1824
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2206 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTGGGCGATCTCCT 2265

Qy  1825 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1884
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2266 AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGATGAGCTGGA 2325

Qy  1885 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 1944
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2326 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 2385

Qy  1945 CATTGACTCTTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAAATTGCGCC 2004
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2386 CATTGACTCTTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAAATTGCGCC 2445

Qy  2005 TCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2064
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2446 TCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2505

Qy  2065 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGAC 2101
         |||||||||||||||||||||||||||||
Db  2506 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGAC 2542
```

RESULT 13
AX538621

Query Match   47.4%; Score 996; DB 6; Length 5462;
Best Local Similarity 100.0%; Pred. No. 1.3e-225;
Matches 996; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1106 CATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAA 1165
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1670 CATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAA 1729

Qy  1166 GCTGAAACAACTGCCAATGTCCTACGAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGAC 1225
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1730 GCTGAAACAACTGCCAATGTCCTACAGATGTCCTACAGATGTCCTAGAGAAGAC 1789

Qy  1226 TCCAAGGGAGTAAAAGAGCTGATGAACACAATGGCAAGACCTCCAAGGTGAAATTGAAGCT 1285
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1790 TCCAAGGGAGTAAAAGAGCTGATGAACACAATGGCAAGACCTCCAAGGTGAAATTGAAGCT 1849

Qy  1286 CACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAAATCTGAGATCCCTGGAA 1345
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1850 CACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAAATCCTGAGATCCCTGGAA 1909

Qy  1346 GGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTCGAGT 1405
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1910 GGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTCGAGT 1969

Qy  1406 GAACTTCGAAAAAGTCTTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGG 1465
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1970 GAACTTCGAAAAAGTCTTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGG 2029

Qy  1466 AAGCGTCTGCACCTTTCTCTGCAGGACACTTCGTGGTGGCTACAGTTGAAAGATGATGAA 1525
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2030 AAGCGTCTGCACCTTTCTCTGCAGGACACTTCGTGGTGGCTACAGTTGAAAGATGATGAA 2089

Qy  1526 TTAAGCCGCAGGCACCTATTGGAGCGGGACTTTCCAGCAGTTCAGAAGCAGCAGCAGCATGTA 1585
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2090 TTAAGCCGCAGGCACCTATTGGAGCGGGACTTTCCAGCAGTTCAGAAGCAGAACGATGTA 2149

Qy  1586 CATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAG 1645
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2150 CATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAG 2209

Qy  1646 ACTGTACGAGAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGAG 1705
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2210 ACTGTACGAGAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGAG 2269

Qy  1706 CCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAAGTCACTCGGCTTCACGAAAGCAG 1765
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2270 CCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCACTACGAAAGCAG 2329

Qy  1766 GCTGAGGAGGTCAATACTGAGTGGGAAAAATGAACCTGCACTCCGTGACTGGCGAGAGA 1825
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2330 GCTGAGGAGGTCAATACTGAGTGGGAAAAATGAACCTGCACTCCGTGACTGGCGAGAGA 2389

Qy  1826 AAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGACGAGCTGGAC 1885
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2390 AAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGACGAGCTGGAC 2449

Qy  1886 CTCAAGCTGCGCCAAGCTGAGTGTGATCAAGGGATCCTGGGCAGCCCGTGGGCCGATCTCCTC 1945
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2450 CTCAAGCTGCGCCAAGCTGAGTGTGATCAAGGGATCCTGGGCAGCCCGTGGGCCGATCTCCTC 2509

Qy  1946 AATTGACTCTTCTCCAAGATCACCTCGAAGATCAAGGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCT 2005
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2510 AATTGACTCTTCTCCAAGATCACCTCGAAGATCAAGGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCT 2569

Qy  2006 CTGAAAGAGACGTGAGCCACGTCAATGACCTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATT 2065
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2570 CTGAAAGAGACGTGAGCCACGTCAATGACCTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATT 2629

Qy  2066 CAGCTCTCACCGTATAACCTCAGCACTCTGGAAGAC 2101
         ||||||||||||||||||||||||||||||||
Db  2630 CAGCTCTCACCGTATAACCTCAGCACTCTGGAAGAC 2665
```

RESULT 14
AF070485

```
CCESSION     AF070485
ERSION       AF070485.1  GI:3982750
EYWORDS      .
OURCE        Canis familiaris (dog)
 ORGANISM    Canis familiaris
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
EFERENCE     1  (bases 1 to 13887)
 AUTHORS     Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
 TITLE       Direct Submission
 JOURNAL     Submitted (04-JUN-1998) ANRI, Pathology, University of Western
             Australia, Verdun Street, Nedlands, WA 6018, Australia
EATURES      Location/Qualifiers
     source  1..13887
             /organism="Canis familiaris"
             /mol_type="mRNA"
             /strain="breed Golden retriever"
             /db_xref="taxon:9615"
     5'UTR   <1..289
     CDS     290..11332
             /codon_start=1
             /product="dystrophin"
             /protein_id="AAC83646.1"
             /db_xref="GI:3982751"
             /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
             QDGRRLLDLLEGLTGQKLPKEGSTRVHALNNVNKALRVLQNNVDLVNIGSTDIVDG
             NHKLTLGLIWNIILHWQVQNVMGNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
             TSWSDGLALNALIHSRPDLFDWNSVVCQQSAEVEMLPRPSQVTREEHFQIHHQMYSQ
             ATTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPSQHLETPEDKSFGRS
             QITVSLAQGYERAPSFPKPRFKSYAYTQAAYVTSDPTRSPLPSQHLETPEDKSFGRS
             LTETEANLDSYQTALEEVLSWLLSAEDALQAQEISNDVEEVKEQFHTEGVMMDLTS
             HQGRVGNVLQLGSQLIGTGKLSEDEETEVQEQMNLLNSRWECLRVASMEKQSNLHKVL
             MDLQNQQKELNDLWDLTKTEERTRKMEKEPLGPIEDLKRQVQQHKVLQEDLEQEQVRV
             NSLTHMVVVDESSGDHATPAALEEQLKVLGDRWANICRWTEDRWVLLQDILLKWQRFT
             EEQCLFSAWLSEWKSEKDAVNKIHTTGFKQSEVLSNLQKLAVLKTDLEKKKQTMDKLCSL
             NQDLLSALKNTVVAHKMEAWLDNFAQRWDNLVQKLEKSSAQISQAVTTQPSLTQTTV
             METVMVTMVTREHILVKHAQEELPPPQKKRQIIVDSEIRKRLDVDITELHSWITRSE
             AVLQSPEFAIYRKSGNFSDLKEVNAIEREKAEKFRKLQDASRSAQLVEQMVNEGVN
             ADSIKQASEOLNSRWIEFCQLLSERLNWLEYQNNIITFYNQLQQLEQMTTAENWLKT
             QPTTTSEPTAIKSQLKICKDEINRLSALQPIERLKIQSIALKEKGQGPMFLDADFVA
             FTNHPNQVFADVQAREKELQTIFDSLPPMRYQETMSTILTWIQQSETKLSIPQVTVTE
             YDIMEQRLGELQALQSSLOEQQNGLNYLSTTVKEMSKKAPLSDISRKYQSEPEEIEGR
             WKKLSSQLVEHCQKLEEQMAKLRKIQNHIKTLKKWITEVDVFLKEEWPALGDSEILKR
             QLKQCRLLVNDIQTIQPSLNSVNEGAQKMKNEAEPEFAGRIESTELRELNTQWDYMCRQ
             VYARKEALKGGLDKTVSLQKDLSEMHEWMTQAEEYLERDFEYKTPDELQTAVEEMKR
             AKEEAQQKEAKVKLLTESVNSVIAQAPPAAQALKKELDTLTTNYQWLCTRLNGKCKT
             LEEVWACWHELLSYLEKANKWLSEVEVKLKTTENISGGAEEIAEVLDSLENLMQHSED
             NPNQIRILAQTLTDGGVWDELINELETFNSRWRELHEEAVRQRLLEQSIQSAQEIE
             KSLHLIQESLSSIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLEEMKKHNQGK
             ETAQRVLSQIDVAQKKLQDVSMKFRLFQKPANFEQRLQESKMILDEVKMHLPALETKS
             VEQEVVQSQLNHCVNLYKSLSEVSEVEMIKTGRQIVQKKQTENPKBLDERVTALKL
             HYNELGAKVTERKQQLEKCLKLSRKMREMNALTEWLAATDMELTKRSAVEGMPSNLD
             SEVAWGKATQKEIEKQKVHLKSVTEVGSALKTVLGKKEMVLEDKLSLLNSNWIAVTSR
             AEEWLNLLLEYQKEMETFDQNVDYITNWIIQADALLDESEKKKPQQKEDILKRLKAEM
             NDIRPKVDSTRDQAANLMANRGDHCRKVVEPKISELNHRFAAISHRIKTGKASIPLKE
             LEQFNSDIQKLLEPLEAEIQQGVNLKEEDFNKDMSEDNEGTVKELLQRGDNLQQRITD
             ERKREBIKIKQQLLQTKHNALKDLRSQRRKKALEISHOWYQYKRQADDLLKCLDDIEK
             KLASLPEPRDERKIKEIDRELQKKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRW
             REISKFAQPRRLNFAQITHVEESVVAMTEDMPLEISYVPSTYLTEITHVSQALSEV
             EELLNAPDLCAQDFEDLFKQEESLKNIKDSLQQISGRIDIHNKKTAALHSATPAERA
             KLQEALSRLDFQWERVNMYKDRQGRFDRSVEKWRRFHYDMKILNQMLTEAEQFLKKT
             QIPENEWEAHKYKWYLKELQDGIGQRQSVRVNATGEBIIQQSSKTDASILQEKLGSL
             NLRWQEVCKQLAERKKRLLEQKNILSEFQRDVNEFVLWLSEAADNVANIPLEPGNEQQL
             KEKLEQVKEIRAHVKDLGQLEQIKQLNELGVLLLWLSPIRNQLEIYNQPNQTGPFDIKEIB
             VSRNLPKEIEAHVKDLGQLEDIIKQKATLQEQRRPQLELITAAQNLKNKTSN
             VAVQAKQPDVEGILSKGQHLYKEKPATQPAKRKLEDLSSDWKVVTQLLQELRAKQPGP
             APGLTTVRAPPSQTVTLVTQPAVTKETAISKLEMPSSILLEVPALADFNRAWTELTDW
             LSLLDRVIKSQRVMVGLEDLINEIKQKATIQLDEQRRPQLELITAAQNLKNKTSN
             QEARTIITDRIERIQSQWDEVQEHLQNRRLQLTEMLKDSTQWLEAKBEAEQVLGQARA
             KLESWKEAPVTVDAIQKKITETKQLAKDLRQWINDVDANDLALKLLRDYSADDTRKV
             HMITENIMASIHKELSREAALEETHRLLQQFPLDLEKFLAWLTEATTANVLQD
             ATHKERLLEDSKGVRELMKQWQDLQGEIEAHTDIYHNLDENGQKVLRSLEGSDDAALL
             QRRLDMNFKWSELRKKSLNIRSHLEASDQWKRLHLSLQELLVWLQLKDDELSRQAP
```

Right column (continuation of translation):

```
             IGGDFPAVQKQNDVHRAPKRELKTKEPVIMSLTETVRIFLTEQPLEGLEKLYQEPREL
             PPERAQNVTRLLRKQAREVNTQWEKLNVHSADWQRKIDEALERLQELQEATDELDLK
             LRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEITPLKENVSYVNDLARQLTTLGI
             QLSPYNLNTLEDLNTRWKLLQVAIEDRIRQLHEAHRDFGPASQHFLSTSVQGPWERAI
             SPNKVPYYINHETQTTCWDHPKVTELYQSLADLNNVRPSAYRTAMKLRRLQKALCLDLL
             LLNVDTGRTGRIRVLSFKTGISLCKAHLEDKYRVLFKQVASSTGDCDQRRLGLLH
             DSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIBAALFLDWMRLEPQSMVWLPV
             LHRVAAAETAKIQAKCNICKECPIIGFRYRSLKHFNYDICQSCFFSGRVAKGHKMHYP
             MVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLEGDNMETPVTLI
             NFWFVDSAPASSPQLSHDDTHSRIEHYASRLKKMENSNGSYLNDSISPNESIDDEHLL
             IQHYWRSLNQESPSLQPRSPAQILISLSEERGELERILADLEGRNNLQAEYDRLKQ
             QEHEKGLSPLPSPPEMMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDHNKQLE
             SQLRLRQLLEQPQAEAKVNGTTVSPPSTSLQRSDSSQPMLRVVGSQTSESSMGEEDL
             LSPPQDTSTGLEEVMEQLNHSFPSSRGRNTPGKPMREDTM"
     3'UTR   11333..>13887
repeat_region  11789..11828
             /rpt_type=tandem
             /rpt_unit="ca"

ORIGIN

Query Match         44.2%;  Score 929;  DB 4;  Length 13887;
Best Local Similarity   90.0%;  Pred. No. 1.1e-209;
Matches  995;  Conservative  0;  Mismatches 110;  Indels  0;  Gaps  0;

Qy      1 GAGCCATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
          |||| ||||| |||||||||| ||||||||||||  ||||| ||||| |||||||||||
Db   1183 GAGCTATGCCTACACACAGGCGTGCTTATGTCACCCACTTCTGACCCCACACGGAGCCCACT 1242

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  120
          |||||||||||||||||||| |||||| |||||||| ||||||||||||||||||||||
Db   1243 TCCTTCACAGCATTGGAAACTCCTGAAGACAAGTCATTGGCCGGTCATTGACAGAGAC 1302

Qy    121 TGAAGTAAAACCTGGACCGTTATCAAACACAGCTTTAGAAGAAGTATTATCGTGGCTCTTTC  180
          |||| ||||||||||||| |||||||||||||||||||||||||||||| |||||| |||||
Db   1303 CGAAGCAAACCCTGGACAGTTATCAAACAGCTTTGGAAGAAGTACTCTCGTGGGCTCTTTTC 1362

Qy    181 TGCTGAGGACACATTGCAAGCACTGCAAGCCCCAAGGAGAGATTCTAATGATGCGGAAAGA  240
          ||| || |||| ||||| ||||||||||||||||||||||||||||||||| ||||| |||||
Db   1363 AGCTGAGGATGCACTGCAAGCCCCAAGGAGAGATTCTAATGATGTGGAAGAAGTGAAAGA 1422

Qy    241 CCAGTTCATACTCATGAGGGGTACATGAAGTTGACAGCCCATGCAGAGTTGGTGG  300
          |  ||| ||||| |||||||| |||||| |||||||||| ||||| ||||||| || |||||
Db   1423 ACAATTTCATACTCATGAGGGGATATATGAAGTTGACACTTGACATCCCATGCAGGACACGGTTCGG 1482

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAAATTATCAAGATGAAGA  360
          |||||||||||||||||||||||| ||||||||||||||||||||||| |||||||||||||
Db   1483 TAATGTTCTCCAACTGGGAAGTCAACTGATTGGAAACAGGGAAAATTATCAAGATGAAGA 1542

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCAAAATCAAGATGGGAATGCCTCAGGGTAGC  420
          || ||||| || |||||||| ||||||||||||||||||||| ||||||||| |||||||||
Db   1543 AACCGAAGTGCGAGGAACAAATGAATGAATCCTCAATCAAGATGGGAATGCCTCAGGGTAGC 1602

Qy    421 TAGCAATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCCCAGAATCAGAAAACT  480
          |||||| ||||||||||||||||||||| ||||||||||||||||||||||||||| |||||
Db   1603 TAGCAATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCCCAGAATCAGCAACT 1662

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAAAAATGAGGAAGAA  540
          ||||||||||||||||||||||||||||| ||||||| |||||| ||||||||
Db   1663 GAAAGAGTTAAATGACTGGCTAACCAAACAGAGAACAACAAGGAAAAATGAGGAAGGA 1722

Qy    541 GCCTCTTGGACCTGATCTTGAAGACAATTGGCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  600
          ||||| ||||||| ||||||||||||||||||| ||||||| |||||||||| |||||||||||||
Db   1723 GCCCCTTGGACCTGATATTGAAGACAATTGGCTTAAAACGCCTAAAATACAACAACAACATAAGTGCTTCA 1782

Qy    601 AGAAGATCTAGAACAAGAACCAAGTCAGGGTCAATTCTCTCACTCCACATGGGTGGTAGT  660
          ||||||||||||||||||||||||||||| |||||| ||||||||| |||||||| ||||||
Db   1783 AGAAGACTTAGAACAAGAACAACAAGTCAGGGTCAATTCCCCACTCTCATATGGGTGGTAGT 1842

Qy    661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACCACCAGTTAAGGTATTGGG  720
          |||| ||| |||| ||||||||| |||| |||||||| |||||||||||||| |||||||||||
Db   1843 CGATGAATCAGTGGAGAACCAGCCATGCAACTGCTGCTTTGGAAGAACAACCAGTTAAGGTACTGGG 1902
```

```
Qy   721 AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1903 AGATCGATGGGCAAACATCTGTAGGTGACAGAAGATCGCTGGGTTCTTTTACAAGACAT 1962

Qy   781 CCTTCTCAAATGGCAAGTCTTACTGAAGACAGTCGCTTTTGTGCATGGCTTTCGA 840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1963 CCTTCTAAATGGCAGGTCTTTACTGAAGACAGTGCCTTTTGTGCATGGCTTTCGGA 2022

Qy   841 AAAGAGAGTGAACAAGAGATTCACCACAACTGGCTTAAAGATCAAAATGAAATGTT 900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2023 GAAGGAAGATGCAGTGAGTGAACAAGATTCACCACAACTGGCTTAAGGATGAAGTGTT 2082

Qy   901 ATCAAGTCTTCAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2083 ATCAAATCTTCAGAACTGCGTCGTCTTAAAAACAGATCTGGAAAAGGAAGAGCAAACCAT 2142

Qy   961 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
             ||||||||||||||||||||||||||||||||||||||||||||
Db  2143 GGACAAACTCTGCCTCACTCAACCAAGACCTTCTTTCAGCGCTGAAATTGGGATAATTTAGTCCAAAA 2202

Qy  1021 CCAGAGACGGAAGCCATGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTTCAAAA 1080
             |||||||||||||||||||||||||||||||||
Db  2203 CCACAAGATGGGAAGCCATGCTGGACAACTTTGCCCAGCGCTGGGATAATTAGTTCCAGAA 2262

Qy  1081 ACTTGAAAAGAGTACAGCACAACACT 1105
             |||||||||||||||||||
Db  2263 ACTTGAAAAAGTTCAGCACCACAGATT 2287
```

```
RESULT 15
E30220
LOCUS        E30220                      4402 bp    DNA        linear   PAT 18-JUN-2001
DEFINITION   Shortened dystrophin.
ACCESSION    E30220
VERSION      E30220.1   GI:13017027
KEYWORDS     JP 199318467-A/3.
SOURCE       unidentified
ORGANISM     unidentified
               unclassified.
REFERENCE    1  (bases 1 to 4402)
AUTHORS      Sinichi,T.
TITLE        Shortened dystrophin
JOURNAL      Patent: JP 199318467-A 3 24-NOV-1999;
             SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY
COMMENT      OS   Unidentified
             PN   JP 199318467-A/3
             PD   24-NOV-1999
             PF   08-MAY-1998 JP 1998142134
             PI   SINICHI TAKEDA
             PC   C12N15/09,A61K48/00,C12N15/00
             CC   Strandedness: Both;
             CC   Topology: Linear;
             FH   Key          Location/Qualifiers
             FT   source       1..4402
                               /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1..4402
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN
```

```
Query Match            43.4%;   Score 911.6;   DB 6;   Length 4402;
Best Local Similarity  99.6%;   Pred. No. 1.5e-205;
Matches  914;   Conservative  0;   Mismatches  4;   Indels  0;   Gaps  0;

Qy     1 GAGCTATGCCTACACACAGCCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1095 GAGCTATGCCTACACACAGCCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1154

Qy    61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1155 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1214
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAGTATTATCGTGGCTTCTTTC 180

Db  1215 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAGTATTATCGTGGCTTCTTTC 1274
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATCATGTGGAAGTGGTGAAAGA 240

Db  1275 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATCATGTGGAAGTGGTGAAAGA 1334
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241 CCAGTTCATATCTCATGAGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 300

Db  1335 CCAGTTCATCTCATGAGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGGT 1394
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360

Db  1395 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1454
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361 AACTGAGTACAAGACAGAGATGAATCCCTAAATTCCTAAATTCAAGATGGGAATGGGAATGCCTCAGGTAGC 420

Db  1455 AACTGAGTACAAGAGACAGAGATGAATCCCTAAATTCCTAAATTCAAGATGGGAATGGGAATGCCTCAGGGTAGC 1514
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 480

Db  1515 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAATGATCTCCAGAATCAGAAACT 1574
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA 540

Db  1575 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA 1634
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 GCCTCTTGGACCTGATCTTGAAGACTTAAAACGCCAAGTACAACACATAAGGTGCTTCA 600

Db  1635 GCCTCTTGGACCTGATCTTGAAGACTTAAAACGCCAAGTACAACACATAAGGTGCTTCA 1694
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601 AGAAGATCTAGACAAGACAAGTCAGGTCAATTCTCTCACTGGTGGTGGTGGTAGT 660

Db  1695 AGAAGATCTAGACAAGAACAAGTCAGGGTCAATTCTCTCACTGGTGGTGGTGGTAGT 1754
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 TGATGAATCTAGTGGACATCACGCAACTCGCTGCTTTGGAAGAACACTTAAGGTATTGGG 720

Db  1755 TGATGAATCTAGTGGAGATCACGCAACTCGCTGCTTTGGAAGAACACTTAAGGTATTGGG 1814
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721 AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 780

Db  1815 AGATCGATGGGCAAACATCTGTAGGTGACAGAAGATCGCTGGGTTCTTTTACAAGACAT 1874
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTTAGTGCATGGCTTTCAGA 840

Db  1875 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1934
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841 AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 900

Db  1935 AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1994
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901 ATCAAGTCTTCAAAAACT 918

Db  1995 ATCAAGTTCTTCGAGAAAGT 2012
             ||||||||     ||||
```

Search completed: April 5, 2004, 02:39:35
Job time : 5504.46 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

un on: April 4, 2004, 16:52:43 ; Search time 2799.18 Seconds
(without alignments)
17079.796 Million cell updates/sec

itle: US-09-845-416-14_COPY_900_2500
erfect score: 1601
equence: 1 gagctatgcctacacacagg.......taccagtcttttagctgacct 1601

coring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

earched: 27513289 seqs, 14931090276 residues

otal number of hits satisfying chosen parameters: 55026578

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	Query %	Match Length	DB	ID	Description
1	1091	68.1	3870	11	BC036103	BC036103 Homo sapi
2	854.2	53.4	3056	11	AK044536	AK044536 Mus muscu
3	833.8	52.1	5691	29	AY399453	AY399453 Homo sapi
4	621	38.8	5697	29	AY399455	AY399455 Mus muscu
5	585.2	36.6	728	14	CB228986	CB228986 AGENCOURT
6	502	31.4	5676	29	AY399454	AY399454 Pan trogl
7 c	497	31.0	595	14	CB177816	CB177816 is21c01.x
8	481.2	30.1	1047	14	CB850319	CB850319 MRA-0070
9	332.6	20.8	2334	11	BC011062	BC011062 Mus muscu
10	322.6	20.1	9915	29	AY407022	AY407022 Homo sapi
11	301.8	18.9	9691	29	AY407024	AY407024 Mus muscu
12	299.8	18.7	9096	29	AY407023	AY407023 Pan trogl
13	264.8	16.5	1122	14	CF109978	CF109978 Shultzomi
14	255.4	16.0	772	12	BI250598	BI250598 60299365 9
15	237.4	14.8	3753	11	AK081426	AK081426 Mus muscu
16	223.8	14.0	826	14	CD656896	CD656896 AGENCOURT
17	220.6	13.8	696	12	BJ075057	BJ075057 BJ075057
18	216.4	13.5	784	12	BG212445	BG212445 RST32032
19	208.6	13.0	630	9	AL855376	AL855376 AL855376
20	205.2	12.8	851	13	BU201022	BU201022 603952191
21	202.4	12.6	502	9	AL602076	AL602076 DKFZp313B
22	201	12.6	2874	29	AY408546	AY408546 Homo sapi
23	199.4	12.5	2874	29	AY408548	AY408548 Mus muscu
24	171	10.7	479	13	BQ304046	BQ304046 QV2-BT063
25	169.8	10.6	835	12	BI553820	BI553820 603190772
26	169.2	10.6	732	13	BU107880	BU107880 603109863
27 c	166.6	10.4	646	12	BI289102	BI289102 UI-R-DKO-
28	166.2	10.4	907	13	BU122401	BU122401 603003073
29	165.4	10.3	681	13	BU301653	BU301653 603609005
30	158.6	9.9	615	14	CA377239	CA377239 655768 NC
31 c	158.2	9.9	569	29	CE228097	CE228097 tigr-gss-
32 c	157.2	9.8	427	10	BB817979	BB817979 CM2-BN027
33	154.6	9.7	405	14	CB811308	CB811308 AWGNNUC:S
34	154.4	9.7	644	13	BU313510	BU313510 603540290
35	152.2	9.5	1541	11	AK034383	AK034383 Mus muscu
36	151.2	9.4	3051	11	BC036095	BC036095 Homo sapi
37 c	150.4	9.4	515	28	AZ780914	AZ780914 2M0018010
38	149.6	9.3	697	13	BU456556	BU456556 603777141
39	146.2	9.1	2874	29	AY408547	AY408547 Pan trogl
40	144.8	9.0	434	13	BQ375536	BQ375536 PM1-TN012
41	141	8.8	763	14	CB518960	CB518960 UI-M-GH0-
42 c	136.4	8.5	466	14	CD549993	CD549993 B0305E01-
43	131.6	8.2	410	13	BQ349936	BQ349936 PM1-HT034
44	131.6	8.2	778	14	CB524596	CB524596 UI-M-FY0-
45	130.2	8.1	426	13	BQ319056	BQ319056 IL5-CT051

ALIGNMENTS

RESULT 1
BC036103
LOCUS   BC036103   3870 bp   mRNA   linear   HTC   19-NOV-2003
DEFINITION Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker
            types), mRNA (cDNA clone IMAGE:5274415), with apparent retained
            intron.
ACCESSION   BC036103
VERSION   BC036103.1   GI:22271310
KEYWORDS   HTC.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3870)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE 2 (bases 1 to 3870)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (31-JUL-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
          cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
          Toshiyuki and Piero Carninci (RIKEN)
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Sequencing Group at the Stanford Human Genome
          Center, Stanford University School of Medicine, Stanford, CA 94305
          Web site: http://www-shgc.stanford.edu
          Contact: (Dickson, Mark) mcd@paxil.stanford.edu
          Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
          R. M.

          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 48 Row: f Column: 7
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Similarity but not
          identity to protein
          This clone has the following problem: retained intron.
FEATURES Location/Qualifiers
     source 1..3870
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="IMAGE:5274415"
          /tissue_type="Brain, hippocampus"
          /clone_lib="NIH_MGC_95"
          /lab_host="DH10B"
          /note="Vector: pBluescript"
ORIGIN

Query Match   68.1%; Score 1091; DB 11; Length 3870;
Best Local Similarity   99.9%; Pred. No. 9.7e-237;
Matches 1102; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCCTCTGACCCTACACGGAGCCCATT    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 GAGCTATGCCTACACACAGGCTGCTATGTCACCACCCTCTGACCCTACACGGAGCCCATT  1102

Qy     61 TCCTTCACAGCAGCTTGGAAGCTCCTGAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 TCCTTCACAGCAGCTTGGAAGCTCCTGAGACAAGCAAGTCATTGGCAGTTCATTGATGGAGAG  1162

Qy    121 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  1222

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  1282

Qy    241 CCAGTTCATACTCATGATGATGATGGTACATGATGGGTACAGAGCCCATCAGGGCCGGGTGG   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1283 CCAGTTTCATACTCATGATGATGGTACATGATGGGTACAGCCCCATCAGGGCCGGGTTGG  1342

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTAGTGATTGGAACAGGAAAATTATCAGAGATGAAGA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1343 TAATATTCTACAATTGGGAAGTAAGCTAGTGATTGGAACAGGAAAATTATCAGAGATGAAGA  1402

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCAAATTCAAGATGGGAATGGCCTCAGGGTAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1403 AACTGAAGTACAAGAGCAGATGAATCTCCAAATTCAAGATGGGAATGGCCTCAGGGTAGC  1462

Qy    421 TAGCATGGAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATC-GAAACT   479
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   1463 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCCAGAATCAGAAAACT  1522

Qy    480 GAAAGAGTTGATGACTGGCTAACAAAACAGAACAAGGAAAATGGGAGGAAGA   539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1523 GAAAGAGTTGATGACTGGCTAACAAAACACAGAGAAGGACAAGGAAAATGGGAGGAAGA  1582

Qy    540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA   599
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1583 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  1642

Qy    600 AGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCCACATGGTGGTGGTAGT   659
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1643 AGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCCACATGGTGGTGGTAGT  1702

Qy    660 TGATGAATCTAGTGGAGATCACGCAACTGCTTGGAAGAACAACTGCTTAAGGTATTGGG   719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1703 TGATGAATCTAGTGGAGATCACGCAACTGCTTGGAAGAACAACTGCTTAAGGTATTGGG  1762

Qy    720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT   779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1763 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT  1822

Qy    780 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTGTGCATGGCTTTCAGA   839
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1823 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTGTGCATGGCTTTCAGA  1882

Qy    840 AAAGAGATGCAGTGAACAGAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT   899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1883 AAAGAGATGCAGTGAACAGAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT  1942

Qy    900 ATCAAGTCTTCAAAAACTGGCCGTTTCAAAAACTGGATCGAATCTAGAAAAAGAAAATCCAT   959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1943 ATCAAGTCTTCAAAAACTGGCCGTTTCAAAAACTGGATCGAATCTAGAAAAAGAAAATCCAT  2002

Qy    960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC  1019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2003 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC  2062

Qy   1020 CCAGAAGACGGAAGCATGGCCGGATAACTTTGCCCGGTGTTGGGGATAATTAGTCAAAA  1079
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2063 CCAGAAGACGGAAGCATGGCCGGATAACTTTGCCCGGTGTTGGGGATAATTAGTCAAAA  2122

Qy   1080 ACTTGAAAAGAGTACAGCACAGA  1102
          |||||||||||||||||||||||
Db   2123 ACTTGAAAAGAGTACAGCACAGA  2145
```

RESULT 2
AK044536
LOCUS   AK044536   3056 bp   mRNA   linear   HTC 20-SEP-2003
DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library,
            clone:A930019F21 product:dystrophin, muscular dystrophy, full
            insert sequence.
ACCESSION   AK044536
VERSION   AK044536.1 GI:26090404
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS   Carninci,P. and Hayashizaki,Y.

```
TITLE      High-efficiency full-length cDNA cloning
JOURNAL    Meth. Enzymol. 303, 19-44 (1999)
MEDLINE    99279253
PUBMED     10349636
EFERENCE   2
AUTHORS    Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
           Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE      Normalization and subraction of cap-trapper-selected cDNAs to
           prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL    Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE    20499374
PUBMED     11042159
EFERENCE   3
AUTHORS    Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
           Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
           Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
           Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
           Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
           Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
           Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE      RIKEN integrated sequence analysis (RISA) system-384-format
           sequencing pipeline with 384 multicapillary sequencer
JOURNAL    Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE    20530913
PUBMED     11076861
EFERENCE   4
AUTHORS    The RIKEN Genome Exploration Research Group Phase II Team and the
           FANTOM Consortium.
TITLE      Functional annotation of a full-length mouse cDNA collection
JOURNAL    Nature 409, 685-690 (2001)
EFERENCE   5
AUTHORS    The FANTOM Consortium and the RIKEN Genome Exploration Research
           Group Phase I & II Team.
TITLE      Analysis of the mouse transcriptome based on functional annotation
           of 60,770 full-length cDNAs
JOURNAL    Nature 420, 563-573 (2002)
EFERENCE   6
           (bases 1 to 3056)
AUTHORS    Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
           Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
           Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
           Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kaeukawa,T.,
           Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
           Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
           Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
           Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
           Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
           Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
           Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
           Muramatsu,M. and Hayashizaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
           Physical and Chemical Research (RIKEN), Laboratory for Genome
           Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
           RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
           Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
           URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
           Fax:81-45-503-9216)
OMMENT     cDNA library was prepared and sequenced in Mouse Genome
           Encyclopedia Project of Genome Exploration Research Group in Riken
           Genomic Sciences Center and Genome Science Laboratory in RIKEN.
           Division of Experimental Animal Research in Riken contributed to
           prepare mouse tissues.
           Reina RNA was provided by Dr. Stefano Gustincich (Department of
           Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
           02115, USA) whose assistance is gratefully acknowledged. Please
           visit our web site for further details.
           URL:http://genome.gsc.riken.go.jp/
           URL:http://fanton.gsc.riken.go.jp/.
EATURES    Location/Qualifiers
source     1..3056
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="C57BL/6J"
```

```
                    /db_xref="FANTOM_DB:A930019F21"
                    /db_xref="MGI:2479854"
                    /db_xref="taxon:10090"
                    /clone="A930019F21"
                    /tissue_type="retina"
                    /clone_lib="RIKEN full-length enriched mouse cDNA library"
                    /dev_stage="adult"
misc_feature    1..3056
                    /note="dystrophin, muscular dystrophy (MGD|MGI:94909,
                    GB|NM_007868, evidence: BLASTN, 100%, match=999)"

ORIGIN
Query Match          53.4%;    Score 854.2;  DB 11;  Length 3056;
Best Local Similarity 86.5%;    Pred. No. 5.2e-183;
Matches 954;  Mismatches   0;  Conservative  0;  Indels 148;  Gaps   1;

Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTGACCCTACACGGAGCCCATT    60
          ||||||||||||||||||||||||||||| ||||||||||||||||| ||||||||
Db    944 GAGTTATGCCTTCACACAGGCTGCTTATGTGCCACCTGATTCCACACAGAGCCCCTA  1003

Qy     61 TCCTTCACAGCATTTGGAGCTCTGAGACAAGCAGTCATTTGGCAGTTCATTGATGGAGAG   120
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1004 TCCTTCACAGCATTTGGAGCTCCCAGAGACAAGCAGTCACTTGACAGTTCATTGATGGAGAC  1063

Qy    121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAAGTATTATCGTGGCTCTTTC   180
          |||||||||||| |||||||||||||||||||||||||||||||||| |||||||||
Db   1064 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAGTACTTTCATGGCTTCTTTC  1123

Qy    181 TGCTGAGGACAATTGCAAGCAAGGAGGAGATTCTAATGATGTGGAAGTGGTGAAAGA     240
          || ||||| |||||||||||||||||||||||||| |||||||||||||||| ||||||
Db   1124 TGCCGAGGATACATTGCGAGCACAAGGAGGAGATTCAAATGATGTTGAAGAAGTGAAAGA  1183

Qy    241 CCAGTTCATCATACTCATGAGGGGTACATGGGGTACAGCCCATCAGGGCCGGGTTGG     300
          ||||||| |||||||||||||||| ||||| || ||||||||||||||||||||||||||
Db   1184 ACAGTTTCATGTCTCATGAGGGATTCATGATGGGTCAGACATCTCATCAAGGACTTGTTGG  1243

Qy    301 TAATATTCTACAATTGGAAGTAGCTGATTGGAACACAGAAAAATTATCAGAAGATGAAGA   360
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244 TAATGTTCTACAGTTGGAAGTAGCTCAACTAGTTGAAAAGGAAAATTATCAGAAGATGAAGA  1303

Qy    361 AACTGAAGTACAAGAGCCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCCAGGGTAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAAGATGGGAATGTCTCCAGGGTAGC  1363

Qy    421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT    479
          |||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||
Db   1364 TAGCATGGAAAAACAAAGCAAATCACAAAGTTCTAATGGATCTCCAGAATCAGAAATT  1423

Qy    480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGA    539
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   1424 AAAAGAACTAGATGACTGGTTAACAAAAAACTGAAGAGAGAACTAAGAAAATGGAGGAGA  1483

Qy    540 GCCTCTTGGACCTGATCTTGAAGACATAAGGTGCTTCA    599
          ||||||||||||||||||||||||| |||||||||||
Db   1484 GCCCTTTGGACCTGATCTTGAAGATCTAAAATGCCAAGTACAACAACATAAGGTGCTTCA  1543

Qy    600 AGAAGATCTAGAACAAGAACAAGTCAGGTGCTCAATTCTCTCACTCACATGGTGGTGGTAGT    659
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1544 AGAAGATCTAGAACAGGAGCAGGTCAGGTCAACTCGGTCACTCACATGGTGGTAGTAGTGGT  1603

Qy    660 TGATGAATCTAGTGGAAGATCACGGCAACTGCTGCTTGGAAGAACAACTAAGTATTGGG   719
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1604 TGATGAATCCAGCGCGTGATCTTGAAGATCTCAAAATGCCAAGACAACTTAAGGTACTGGG  1663

Qy    720 AGATCGATGGGCAAACATCCTGTAGATGGACAGCAGAAGACGCGGGTTGTCTTTACAAGACAT   779
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   1664 AGATCGATGGGCAAATATCTCGATGACTGGGACTGAAGACCGCTGAAGGTGTTTTACAAGATAT  1723

Qy    780 CCTTCCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA    839
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   1724 TCTTCAAAATGGCAGCATTTACTGAAGAACAGTGCCTTTTTTTAGTACACATGGCTTTCAGA  1783

Qy    840 AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAGATCAAAATGAAATGTT    899
```

Db   1784 AAAAGAAGATGCAATGAAGAACATTCAGACACAAGTGGCTTAAAGATCAAAATGAAATGAT 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   405 TGAAGTAAACTGGACCGTTATCAAACAGCTTTAGAGAAGAGTATTATCGTGGCTTCTTTC 464

Qy   900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGATCAAGTGGATCTAGAAAAGAAAAAGCAATCCAT 959
          |||||| ||| ||| |||||| |||||
Qy   181 TGCTGAGGACACATTGCAAGCACAAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240

Db   1844 GTCAAGTCTTCACAAAAATATCTACTTTAAAAATAGATCTAGAAATAGAAAAAGCCAACCAT 1903
          |||||| ||| ||| |||||| |||||
Db   465 TGCTGAGGACACATTGCAAGCACAAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 524

Qy   960 GGGCAAACGTATTCACTCAACAAGATCTCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
          || |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241 CCAGTTTCATCTCATGAGGGTACATGATTGATTGACAGCCCATCAGGGCCGGGTTGG 300

Db   1904 GGAAAAACTAAGTCACTGAATCAGATCTACTTTCGGCACTGAAAAATAAGTCAGTGAC 1963
          ||||||| ||||||||||||| |||||||||||||||||||||||||||||||
Db   525 CCAGTTTCATCTCATGAGGGTACATGATTGATTGACAGCCCATCAGGGCCGGGTTGG 584

Qy   1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTTGGGGATAATTTAGTCCAAAA 1079
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGGAAAAATTATCGAAGATGAAGA 360

Db   1964 TCAAAAGATGGGAATCTGGATGGAAAAACTTTGCACAACGTTGGGACCAATTTAACCCAAAA 2023
          ||||||||||||||||||| .||||||||||||||||||| |||||||
Db   585 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGGAAAAATTATCGAAGATGAAGA 644

Qy   1080 ACTTGAAAAGAGTACAGCACAGA 1102
          |||||||||||||||||||||||
Qy   361 AACTGAAGTACAGAGAGCAGATGAATCTCCTAAATCAAGATGGGAATCCCAGAATC-GAAACT 420

Db   2024 ACTTGAAAAGAGTTCAGCACAAA 2046
          |||||||||||||| |||||||
Db   645 AACTGAAGTACAGAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC 704

RESULT 3
AY399453
LOCUS   AY399453   5691 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY399453
VERSION   AY399453.1 GI:39755442
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   705 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 764

REFERENCE   1 (bases 1 to 5691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.

Qy   480 GAAAGAGTTGAATGACTGGCTAACAGAAAACAGAGAAGAACAGAGAAAATGAGGAGGAAGA 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   765 GAAAGAGTTGAATGACTGGCTAACAGAAAACAGAGAAGAACAGAGAAAATGAGGAGGAAGA 824

TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302

Qy   540 GCCTCTTGGACCTGATCTTGAAGACCTAAGACGGCCAAGTACAACACATAAGGTGCTTCA 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   825 GCCTCTTGGACCTGATCTTGAAGACCTAAGACGGCCAAGTACAACACATAAGGTGCTTCA 884

REFERENCE   2 (bases 1 to 5691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.

Qy   600 AGAAGATCTAGAACAAGTCAGGTCAATTCTCCTCACACGTGGTGGTAGT 659
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   885 AGAAGATCTAGAACAAGTCAGGTCAATTCTCCTCACACGTGGTGGTAGT 944

TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA

Qy   660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   945 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1004

COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.

Qy   720 AGATCGATGGGCAAACATCTGAGATGCAGACGAAGACCGCTGGGTTCTTTACAAGACAT 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1005 AGATCGATGGGCAAACATCTGTAGATGCAGACGAAGACCCGCTGGGTTCTTTACAAGACAT 1064

FEATURES   Location/Qualifiers
     source   1..5691
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

Qy   780 CCTTCTCAAATGCAACGTCTTACTGAAGACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
          ||||||||||||||||||||||||||||||
Db   1065 CCTTCTCAAATGCAACGTCTTACTGAAGACAG----------------------------- 1098

     gene   <1..>5691
                     /locus_tag="HCM0229"

Qy   840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAATGAAAATGTT 899
                                                                          -
Db   1099 --- 1099

ORIGIN

Qy   900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAACTGGATCTAGAAAAGAAAAAGCAATCCAT 959
Db   1099 -----------------------GTTTTAAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 1136

Qy   960 GGGCAAACTGTATTCACTCAACAAGATCTCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
Db   1137 GGGCAAACTGTATTCACTCAACAAGATCTCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1196

Query Match   52.1%; Score 833.8; DB 29; Length 5691;
Best Local Similarity   89.1%; Pred. No. 2.5e-178;
Matches 983; Conservative   0; Mismatches   2; Indels 118; Gaps   3;

Qy   1020 CCAGAAGACGGAAGCAAGCATGGCTGGATAACTTGCCCGGTGTTGGGGATATAATTAGTCCAAAA 1079
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1197 CCAGAAGACCGAAGCAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGGATAATTTAGTCCAAAA 1256

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
Db 294 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 353

Qy   1080 ACTTGAAAAGAGTACAGCACAGA 1102
          |||||||||||||||||||||||
Db   1257 ACTTGAAAAGAGTACAGCACAGA 1279

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGGAGAG 120
        ||||||||||||
Db 354 TCCTTCACAGC-------TCCCTGAAGACAAGTCATTGGCAGTTCATTGATGGGAGAG 404

RESULT 4
AY399455
LOCUS   AY399455   5697 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.

Qy 121 TGAAGTAAACTGGACCGTTATCAAACAGCTTTAGAGAAGAGTATTATCGTGGCTTCTTTC 180

CCESSION   AY399455
ERSION   AY399455.1   GI:39755444
EYWORDS   GSS.
OURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
EFERENCE   1 (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
EFERENCE   2 (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
 OMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
EATURES   Location/Qualifiers
     source   1..5697
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10090"
     gene   <1..>5697
                    /locus_tag="HCM0229"
 RIGIN

Query Match   38.8%;   Score 621;   DB 29;   Length 5697;
Best Local Similarity 77.1%;   Pred. No. 5e-130;
Matches 850;   Conservative 0;   Mismatches 135;   Indels 118;   Gaps 3;

1   GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
    ||||||||||||||||||||| ||||||||||| |||| ||| ||||| |||||||||
300 GAGTTATGCCTTCACACAGAGCTGCTTATGTTGGCACCCTGATTCCACACAGAGCCCCTA 359

61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 120
    |||||||||||| ||||||||||||||||||||||||||||||||||||||
360 TCCTTCACAG--------GCTCCCAGAGACAAGTCACTTGACAGTTCATTGATGGAGAC 410

121 TGAAGTAAACCTGACCGGTTATCAAACAGCTTAGAGAAGTATTATCGTGCGTTCTTTTC 180
    || ||||||||||||||||||| ||||||| ||||||||||||| |||||||||||||
411 GGAAGTAAATCTGATAGTTACCAAACAAATGAATCTCTAAATTCCTAAATTCAAGATGAAGTC 470

181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
    |||||||||||| ||||||| |||||||||||||||||| |||||||||||||||||
471 TGCCCAGGATACATTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAAGA 530

241 CCAGTTTCATCATCATGAGGGGTACATGATGATTGACAGCCATCAGGCCGGGTTGG 300
    |||||||||||| |||||||||| ||||||||||||||||||||||||||||||||
531 ACAGTTTCATGCTCATGAGGGATTCATGATGATCTGACATCTCATCAAGGACTTGTTGG 590

301 TAATATTCTACACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
    |||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
591 TAATGTTCTCACAGTTTGGAAGTAAGTCAACTAGTTGGAATCTCAACAAAGGGAAATTATCAGAAGATGAAGA 650

361 AACTGAAGTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
    ||||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||
421 AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCCTAAATTCAAGATGGGAATGTCTCAGGGTAGC 710

421 TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAATC-GAAACT 479
    ||||||||||||| |||||| |||||||| ||| | |||||||||||||| ||||||
711 TAGCATGGAAAAACAAAGCAAATTTACACCAAAAGTTCTAATGGATCTCCAGAATCAGAAATT 770

480 GGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAACAGAAGAACAAGAAAATGGAGGAAGA 539
    |||| | ||||||||||||||| |||||| ||||| | | ||||||||| | ||||||||

771 AAAAGAACTAGATGACTGGTTAACAAAAACTGAAGAGAGAGAGAACTAAGAAAATGAGGAAGA 830
    ||||||||||||||||| ||||||||||||| |||||||||||| ||||||||||||||
540 GCCTCTTGGACCTGATCTTGAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 599

831 GCCCTTGGACCTGATCTTGAAGATCTAAAATCTAAAATCTAAAACACAACATAAGGTGCTTCA 890
    ||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||
600 AGAAGATCTAGAACAAGAACAAGACAAGTCAGGTCAATTCTCTCACTCACCACATGGTGGTGTAGT 659

891 AGAAGATCTAGAACAGAGCAGGTCAGGTCAACTCGCTCACTCCACCTCACTCCACATGTAGTAGTGGT 950
    ||||||||||||||| ||| ||||||||||||| | || ||| ||| ||||| |||
660 TGATGAATCAGTGGAGATCACGCAACTGCTGTTTGAAGAACAACTTAAGGTATTGGG 719

951 TGATGAATCCAGGGTGATCATCGCAACAGCTGCTTTGGAAGAACAACTTAAGGTACTGGG 1010
    ||||||||| ||| | ||||||||||| ||||| ||||||||||||||||||||||||
720 AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGTTCTTTACAAGACAT 779

1011 AGATCGATGGGCAAATATCTGCAGATGGACTGAGGACCTGAGATTGTTTACAAGATAT 1070
    ||||||||||||||| ||||| |||||| ||||| ||| |||| ||||||||||
780 CCTTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTTTAGTGCATGGCTTTTTCAGA 839

1071 TCTTCTAAAATGGCAGCATTTTACTGAAGAACAG--------------------- 1104
    ||||| ||||| ||||||||||| |||||||||
840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 899

1105 --- 1104

900 ATCAAGTCTTCAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959

1105 --------ACTTTAAAAATAGATCTAGAAAAGAAAAAGCCAACCAT 1142

960 GGGCAAACTGTATTCACTCAAACAAGATCTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
    ||||||||||||||||| ||||||||||||||||||||| |||| ||||||||||||
1143 GGAAAAACTAAGTTCACTCAATCAAGATCTACTTCCGGCACTGACTGAAAAAATAAGTCAGTGAC 1202

1020 CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGGTGTGTTGGGATAATTTAGTCCAAAA 1079
    ||||||| |||||| |||||| ||||| ||||||||||||| |||| ||||| |||||
1203 TCAAAAGATGGAAATCTCGGATGTGAAAAACTTTGGACAACGTTGGGACAATTTAACCCAAAA 1262

1080 ACTTGAAAAGAGTACAGCACACAGA 1102
    ||||||| ||||||||||||||||
1263 ACTTGAAAAGAGTTCAGCACAAA 1285

RESULT 5
 CB228986
LOCUS   CB228986   728 bp   mRNA   linear   EST 10-FEB-2003
DEFINITION AGENCOURT_11499247 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
            IMAGE:6884920 5', mRNA sequence.
ACCESSION   CB228986
VERSION   CB228986.1   GI:28280564
KEYWORDS   EST.
SOURCE   Macaca mulatta (rhesus monkey)
  ORGANISM Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1 (bases 1 to 728)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Eliot Spindel
            cDNA Library Preparation: CLONTECH
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3135 row: e   column: 11
            High quality sequence stop: 583.
FEATURES   Location/Qualifiers

ORGANISM Pan troglodytes
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 5676)
  AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
           Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
           gene trios
  JOURNAL Science 302 (5652), 1960-1963 (2003)
  PUBMED   14671302
REFERENCE 2 (bases 1 to 5676)
  AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
           Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
           Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
           them based on alignment.
FEATURES   Location/Qualifiers
     source   1..5676
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene   <1..>5676
                     /locus_tag="HCM0229"
ORIGIN

Query Match   31.4%; Score 502; DB 29; Length 5676;
Best Local Similarity   75.1%; Pred. No. 5.1e-103;
Matches 634; Conservative   0; Mismatches 101; Indels 109; Gaps   2;

Qy   260 GGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGGTGTAATATTCTACAATTGGGA 319
          ||||||||||||||||||| |||||| ||||||||||||||||||| |||| ||||||
Db   529 GGGGTACATGATGATNNNNAGCCCATCAGGGCCGGTTGGTTGGTTAATANNNNNNNAATTGGGA 588

Qy   320 AGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGATGAAGAAACTGAAGTACAAGACAG 379
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   589 AGTNNNNNTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGATGAAGAAACTGAAGTACAAGACAG 648

Qy   380 ATGAATCCTCAAATTCAAGATGGGAATGCCTCAGGGTAGCTGAGTGCTTCAGGGAAAAAAACAAAGC 439
          |||||||| ||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   649 ATGAATCNNNNNATTCAAGATGGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 708

Qy   440 AATTTACATAGAGTTTAATGATCTCCAGAATC-GAAACTGAAAGAGTTGAATGACTGG 498
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   709 NNTTTACATAGAGTTTAATGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGG 768

Qy   499 CTAACAAAAACAGAAGAAGAACAAGGAAAATGAGGAAGAACAAGGAAAATGAGGAAGAGCCCTTGGACCTGATCTT 558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   769 CTAACAAAAACAGAAGAAGAACAAGGAAAATGAGGAAGAACAAGGAAAATGAGGAAGAGCCCTTGGACCTGATCTT 828

Qy   559 GAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCAAGAGATCTCAGAAGACAACAAGAA 618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   829 GAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCAAGAGATTTAGAACAAGAGA 888

Qy   619 CAAGTCAGGTGCATCAATTCTCTCACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGAT 678
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   889 CAAGTCAGGTGCATCAATTCTCTCACTCACATGGTGGTGTAGTTGATGAANNNNANTGGAGAC 948

Qy   679 CACGCAACTGCTGCTTTGGAAGAACAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATC 738
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   949 CACGCAACTGCTGCTTTGGAAGAACAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATC 1008

Qy   739 TGTAGGTGGACAGAAGCCGCTGGGTTCTTTACAAGACATCCTTCTCAAATGGCAACGT 798
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1009 TGTAGGTGGACAGAAGCCGCTGGGTTCTTTACAAGACATCCTTCTCAAATGGCAACGT 1068

Qy   799 CTTACTGAAGAACAGTGCTTTTGTCATGGCTTTCAGAAAAAAGAAGATGCAGTGAAC 858
          |||||||||||||

---

source   1..728
                  /organism="Macaca mulatta"
                  /mol_type="mRNA"
                  /db_xref="taxon:9544"
                  /clone="IMAGE:6884820"
                  /tissue_type="ovary"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NICHD_Rh_Ov1"
                  /note="Organ: ovary;Vector: pDNR-LIB; Site 1: Sfi I;
                  Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
                  Average insert size 1.0-4.0 kb. Tissue pooled from
                  pre-pubertal, post pubertal sn menopausal monkeys.
                  Constructed by Clontech. Note: this is a NICHD Library."
ORIGIN

Query Match   36.6%; Score 585.2; DB 14; Length 728;
Best Local Similarity   96.9%; Pred. No. 4.6e-122;
Matches 618; Conservative   0; Mismatches 18; Indels   2; Gaps   2;

Qy   295 GGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGA 354
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGA 60

Qy   355 TGAAGAAACTGAAGTACAAGAGCAGATGAATCCTCAAATTCAAGATGGGAATGCCTCAG 414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGAAGAAACTGAAGTACAAGAGCAGATGAATCCTCAAATTCAAGATGGGAATGCCTCAG 120

Qy   415 GGTAGCTAGCATGGAAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATC- 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGTAGCTAGCATGGAAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCA 180

Qy   474 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGA 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGA 240

Qy   534 GGAAGAGCCCTTGGACCTGATCTTGAAGAACAGCCAAGTACAACAACATAAGGT 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAAAGAACCCCTTGGACCTGATCTTGAAGAACAGCCAAGTACAACAACATAAGGT 300

Qy   594 GCTTCAAGAGATCTAGAACAAGAACAAGATGCAGTGCAACGTCAGGGTCAATTCTCCACTCACATGGTGGT 653
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCTTCAAGAGATCTAGAACAAGAACAAGATGCAGTGCAACGTCAGGGTCAATTCTCCACTCACATGGTGGT 360

Qy   654 GGTAGTTGATGAATCTAGTGGAGATCACGCTGCTTGGAAGAACAACAACTTAAGGT 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGTAGTTGATGAATCTAGTGGAGATCACGCTGCTTGGAAGAACAACAACTTAAGGT 420

Qy   714 ATTGGGAGATCGATGGGCAAACATCTGTAGATGGGCAACGTCAGGGTTCTTTTACA 773
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   421 ATTGGGAGATCGATGGGCAAACATCTGTAGATGGGCAACGTCAGGGTTCTTTTACA 480

Qy   774 AGACATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCT 833
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AGACATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCT 540

Qy   834 TTCAGAAAAAGAAGATGCAGTGAACAGATTCACACAACT-GGCTTTAAAGAATCAAAATG 892
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   541 TTCAGAAAAAGAAGATGCAGTGAACAGATTCACACAACTGGGCTTTAAGGATCAAAATG 600

Qy   893 AAATGTTATCAAGTCTTCAAAAAACTGGCCGTTTAAAA 930
          |||||||||||||||||||||||||||||||||||||
Db   601 AAATGTTATCAAGTCTTCAAAAAACTGGCCGTATGTAA 638

RESULT 6
AY399454
LOCUS   AY399454   5676 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
           genomic survey sequence.
ACCESSION AY399454
VERSION   AY399454.1 GI:39755443
KEYWORDS   GSS.
SOURCE   Pan troglodytes (chimpanzee)

                                        Metabolism Div. (Alan Permutt Lab), Washington University
                                        School of Medicine, Box 8127, 660 South Euclid Ave., St.
                                        Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                                        314-362-1916, Fax: 314-747-2692."

ORIGIN

     Query Match   31.0%;   Score 497;   DB 14;   Length 595;
     Best Local Similarity 99.8%;   Pred. No. 4.6e-102;
     Matches 508;   Conservative 0;   Mismatches 0;   Indels 1;   Gaps 1;

Qy   1 GAGCTATGCCTACACACCAGGCTGCTTATGTCACCACCCTCGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   509 GAGCTATGCCTACACACCAGGCTGCTTATGTCACCACCCTCGACCCTACACGGAGCCCATT 450

Qy   61 TCCTTCACAGCATTGGAGCTCCTGAGACAGCAAGTCATTGGCAGTTCATTGATGGGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   449 TCCTTCACAGCATTGGAGCTCCTGAGACAGCAAGTCATTGGCAGTTCATTGATGGGGAGAG 390

Qy   121 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAGAAGAAGAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   389 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAGAAGAAGAGTATTATCGTGGCTTCTTTC 330

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGACACACAAGGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   329 TGCTGAGGACACATTGCAAGCACAAGGACACACAAGGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 270

Qy   241 CCAGTTTCATATCTCATGAGGGGTACATGATGGGATTGGACAGCCCATCAGGGCCGGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   269 CCAGTTTCATATCTCATGAGGGGTACATGATGGGATTGGACAGCCCATCAGGGCCCGGGGTTGG 210

Qy   301 TAATATTCTACAATTGGAAGTAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   209 TAATATTCTACAATTGGAAGTAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA 150

Qy   361 AACTGAAGTACAAGAGCACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   149 AACTGAAGTACAAGAGCACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 90

Qy   421 TAGCATGGAAAAACAAAGCAATTACAATTACATAGAGTTTTAATGGATCTCCAGAATC-GAAACT 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   89 TAGCATGGAAAAACAAAGCAATTACAATTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 30

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAAA 508
          |||||||||||||||||||||||||||||
Db   29 GAAAGAGTTGAATGACTGGCTAACAAAAA 1

RESULT 8
CB850319
LOCUS   CB850319   1047 bp   mRNA   linear   EST 01-SEP-2003
DEFINITION MRA-0070 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA sequence.
ACCESSION   CB850319
VERSION   CB850319.1 GI:34380806
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1047)
  AUTHORS   Yu,J., Farjo,R., MacNee,S.P., Baehr,W., Stambolian,D.E. and
            Swaroop,A.
  TITLE   Annotation and analysis of 10,000 expressed sequence tags from
            developing mouse eye and adult retina
  JOURNAL   Genome Biol. 4 (10), R65 (2003)
  MEDLINE   22881944
   PUBMED   14519200
  COMMENT   Contact: Swaroop,A.
            Department of Ophthalmology and Visual Sciences
            Kellogg Eye Center, University of Michigan
            540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
            Tel: 734 615 2246
            Fax: 734 647 0228
            Email: swaroop@umich.edu.

Db   1069 CTTACTGAAGAACAG------------------------------- 1083

Qy   859 AAGATTCACACAACTGGCTTAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTG 918

Db   1084 --- 1083

Qy   919 GCCGTTTTAAAAGCGGATCTAGAAAGAGAAAAAGAAAAAGCAATCCATGGGCAACTGTATTCACTC 978
          ||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db   1084 ---GTTTTAAAAGCGGATNNNGAAAGAGAAAAAGAAAAAGCAATCCATGGGCAACTGTATTNNNNN 1140

Qy   979 AAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCATGG 1038
          ||||||| | || || ||||||||||||||||||||||||||||||||||||||| ||
Db   1141 NNACAAGNTNTTCTNTNAACACTGAAGAATAAGTCAGTGACCCAGAAGACGGAAGCNNGG 1200

Qy   1039 CTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCA 1098
          ||||||||||||||| || ||||||||||| ||| |||||| ||||||||||||||
Db   1201 NNGGATAACTTGCCCGNNNTGGGATAATTNNNCCAAAAACNNGGAAAAGAGTACAGCA 1260

Qy   1099 CAGA 1102
          ||||
Db   1261 CAGA 1264

RESULT 7
CB177816/c
LOCUS   CB177816   595 bp   mRNA   linear   EST 31-JAN-2003
DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6653129 3'
            similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ;, mRNA sequence.
ACCESSION   CB177816
VERSION   CB177816.1 GI:28186206
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 595)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Bistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE   Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
  COMMENT   Other_ESTs: is21c01.y1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 448.
FEATURES   Location/Qualifiers
     source   1..595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6653129"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,

KEYWORDS   HTC.
SOURCE   Mus musculus (house mouse)
   ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 2334)
   AUTHORS   Strausberg,R.
   TITLE   Direct Submission
   JOURNAL   Submitted (25-JUL-2001) National Institutes of Health, Mammalian
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           USA
   REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
           Contact: MGC help desk
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Gilbert Smith, Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Baylor College of Medicine Human Genome
           Sequencing Center
           Center code: BCM-HGSC
           Web site: http://www.hgsc.bcm.tmc.edu/cdna/
           Contact: amg@bcm.tmc.edu
           Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
           Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
           A.N., Gibbs, R.A.
   COMMENT   clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series: IRAK Plate: 23 Row: j Column: 4
           This clone was selected for full length sequencing because it
           passed the following selection criteria: Hexamer frequency ORF
           analysis, similarity but not identity to protein
           This clone has the following problem: retained intron.
FEATURES   Location/Qualifiers
   source   1..2334
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="FVB/N"
           /db_xref="taxon:10090"
           /clone="IMAGE:3979320"
           /tissue_type="Mammary tumor. Metallothionien-TGF alpha
           model. 10 month old virgin mouse. Taken by biopsy."
           /clone_lib="NCI_CGAP_Mam1"
           /lab_host="DH10B"
           /note="Vector: pCMV-SPORT6"
ORIGIN

Query Match   20.8%;   Score 332.6;   DB 11;   Length 2334;
Best Local Similarity   59.0%;   Pred. No. 1.2e-64;
Matches 589;   Conservative 0;   Mismatches 409;   Indels 1;   Gaps 1;

```
Qy    104 AGTTCATTGATGGAGTGAAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTA 163
              ||||||||||||||||||||  ||||| |||||||| ||||||||||||||  ||||||
Db   1042 AGCACCGTCACTGAAGTGGACATGATGTTGGACAGCAGCACCAAAGGAGAGATTGCGCTGAGGAAGTG 1101

Qy    164 TTATCGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGCAAGATGACAT TTCTGATGAT 223
              ||||| ||||||||||||  || |||||||||||   |||||||||||| |||||||||||
Db   1102 CTGACGTGGCTGCTGTCCGCGGAGGACACGTTCCAGGAGCAAGATGTACATGATGATGATGAT 1161

Qy    224 GTGGAAGTGGTGAAAGACCAGTTGCATTCCTCATCATGAGGGGTACATGATGATTTGACAGCC 283
              |||||||||||||||||| ||||||||||| || |||||||||||||  |||||||||||||
Db   1162 GTCGAAGAAGTCAAAGAGCAGCAGTTGCTACCACCATGGAAGTAAGCTAAGCTGATGGAACAGGAAAA 1221

Qy    284 CATCAGGGCCGGTTGGTAATATTCTACAATTGGGAGGATCGATG 343
              || ||||| |||| ||||| |||| ||||||| |||| ||||||
Db   1222 CACCAGAGCAGCGTGGGGAGCGTCCTGCAGGCTGGCAACCAGCTGCATGACACAAGCGGACT 1281

Qy    344 TTATCGAAGATGAAGAAAACTGAAGTACAGAAGAGCAGAGCAGAGAATTCCTAAATTCAAGATGG 403
              |||| |||||||| ||||||| |||||   || ||| | ||||||||||||||||||
Db   1282 CTGTCAGGAGGAGGAGTTTGAGATCCAGGAACAGATGACCTTGCTGAATGCAAGGTGG 1341

Qy    404 GAATGCCTCAGGGTAGCTAGCATGGAGTAGCATGAAGAAAACAAAGCAATTTACATAGAGTTTAATGGAT 463
```

TITLE   Adams,M.D. and Cargill,M.
JOURNAL   Direct Submission
           Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
           Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
           them based on alignment.
FEATURES   Location/Qualifiers
     source   1..9915
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>9915
                     /gene="UTRN"
                     /locus_tag="HCM2752"
ORIGIN

Query Match
Best Local Similarity   20.1%;   Score 322.6;   DB 29;   Length 9915;
Matches 522;   Conservative 0;   Mismatches 314;   Indels 1;   Gaps 1;
                        62.4%;   Pred. No. 3e-62;

Qy 103 CAGTTCATTGATGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTAGAAGAAGT 162
Db 918 CAGCACTGTCACTGAGGTTGACATGGATCTGGACACTATCAGATTGCGTTGAGGGGAGT 977

Qy 163 ATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGA 222
Db 978 GCTGACCTGGTTGCTTTCGTTCTGCTGAGGATGCTTCCAGGAGCAGGATGATATTCTGATGA 1037

Qy 223 TGTGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGGATTTGACAGC 282
Db 1038 TGTTGAAGAAGTCAAAGACCCAGTTGCAACCCATGAAGCTTTATGATGGAACTGACTGC 1097

Qy 283 CCATCAGGCGGCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTAGGTGATTGGAACAGGAAA 342
Db 1098 ACACCAGAGCAGTGTGGGCAGCCGTCCTGCAGGAGGCAACCAACCACTGATCGATAACAAGGAAC 1157

Qy 343 ATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATG 402
Db 1158 TCTGTCAGCAGAAGAAGAATTGAGATTCAGGAACAGATGACCCTGCTGAATGCTAGATG 1217

Qy 403 GGAATGCCTCAGGGTAGCTAGCATGGAAAAAACAAAGCAATTACATAGAGTGTTTAATGA 462
Db 1218 GGAGGCTCTTAGGGTGGGAGTATGGCACAGTCCCGCTGCACGATGTGCTGATGGA 1277

Qy 463 TCTCCAGAA-TCGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAAC 521
Db 1278 ACTGCAGAAGAAGCAACTGCAGCAGCTCTCCGCCTGGTTAACACTCACAGAGACCTAAAACGCCAAGTACA 1337

Qy 522 AAGGAAAATGGAGAAGAGAGCCTCTTGGACCTGATCTTGAAGAACGCCAAGTACA 581
Db 1338 TCGAGAGATGAAACTTGCCCCCTGGATGATGGTGAAAATCTCTACAAAAGCTGCTAGA 1397

Qy 582 ACAACATAAGGTGCTTCAAGAAGTCTAGAACAAGAACAAGTGCAGGGTCAATTCTCTCAC 641
Db 1398 AGAACATAAAAGTTTGCAAAGTGATCTTGAGCGTGAACACAGGTGAAAGTAAATTCACTAAC 1457

Qy 642 TCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGAGTCACGCAACTGCTGCTTTGGAAGA 701
Db 1458 TCACATGGTGTCATTGTGGTGATGAAAACAGTGGTGGAGAGTGCCAAACATCTGTAGATGGCAAA 1517

Qy 702 ACAACTTAAGGTATTGGGAGGATCGATGGGCAAACATCTGTAGATGGCAGAGACCGCTG 761
Db 1518 CCAGTTACAGAAAACTTGGTGAGCGCTGGACAGCGTGATCTTGGACTGAAGAACGCTG 1577

Qy 762 GGTTCTTTACAAGACATCCTTCACGAAATGGCAACGTCTTTACTGAAGAACAGTGCCTTTTT 821
Db 1578 GAATAGGTTACAAGAAATCAATAATAATTGTGGCAGGAATTATTGGAAGAACAGTGCTGTTGTT 1637

Qy 822 TAGTGCATGGCCTTTCAGAAAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAA 881
Db 1638 GAAAGCTTGGTTAACCGAAAAAGAAGAGGCTTTAAATAAAGTCCAGAACAAGCAACTTCAA 1697

Qy 882 AGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTTAAAAGCGGGATCT 938

b 1342 GAGGGCTCCGGGTGGAGAGCATGAGCAGTCCCGGCTCGCACGAGACGCTCGATGGAG 1401
y   464 CTCCGAGATCGA-AACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACA 522

b 1402 CTGCAGAAGAAACAGCTGCAGCAGCTCTCAAGCTGGCTGGCCCTCACAGAGAGCGCCATT 1461
y   523 AGGAAAATGGAGAAGAGAGCCTCTTGACCTGATCTTGAAGAACGCCAAGTACAA 582

o 1462 CAGAAGAATGGAGAATGGAGAGCCTCCCGCTGGGTGATGACCTGCCCTCCCTGCGAGAAGCTGCTTCAA 1521
y   583 CAACATAAGGTGCTTCAAGCAGTCAGGGTCAATTCTCTCACT 642

o 1522 GAACATAAAAGTTTGCAAATGACCTTGAGCTGAACACAGTGAAGGTAAATTCCTTRACT 1581
y   643 CACATGGTGGTGGTAGTTGATGAATCTAGTGGAGAATCACGCAACTGCTGCTTTGGAGAA 702

b 1582 CACATGGTGGTGATTGTGGATGAAAACAGTGGGGAGGAGTGCCACAGTCTTCTCTGGAGAT 1641
y   703 CAACTTAAGGTATTGGGAGGATCGATGCATCGTAGATGGACAAACATCTGTAGATGGCAC 762

b 1642 CAGTTACAGAGAAACTCGGTGAGCGCTGGACAGCGTGATCTGGACTGAAGAACGTTGG 1701
y   763 GTTCTTCTTTACAAGACATCCTTCTCAAATGGCAAGTCTCTACTGAAGAACAGTGCCTTTT 822

b 1702 AACAGGTTGCGAAGAAATCAGTATTCTGTGGCAGGAATTATTGGAAGAGCAGTGTGTCTGTTG 1761
y   823 AGTGCATGGCCTTTCAGAAAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAA 882

b 1762 GAGGGCTTGGGCTTCACCCGAAAAGGAGGAGCCTTTGAATAAAGTTCAAACCAGCAACTTTAAA 1821
y   883 GATCAAAATGAAAGTTTATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAA 942

b 1822 GACCAGAGGAACTAAGTGTCAGTGTCCGCGCTCTGGCTATATTGAAGGAAGACATGGAA 1881
y   943 AAGAAAAAGCAATCCATGGCAGCAACTCTGTATTCACTCAAACAAGATCTCTCTTCAACACTG 1002

b 1882 ATGAAGAGGCAGACTCTGATCAACTCGAGTGAGATTGGCCAGGATGTGGGCAATTACTC 1941
y 1003 AAGAATAAGTCAGTGACCCAGAACGGAGCATGGCTGGATAACTTTGCCCGGTGTTGG 1062

b 1942 AGTAATCCCAGGGCATCTAAGAAGATGAACAGTGACCTCTGAGGAGCTAACACAGAGATGG 2001
y 1063 GATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACAG 1101

b 2002 GATTCTCTGGTTCAGAGACTCGAGACTCTTCTAACCAG 2040

RESULT 10
AY407022
LOCUS   AY407022   9915 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION Homo sapiens UTRN gene, genomic survey sequence.
ACCESSION   AY407022
VERSION   AY407022.1 GI:39762993
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 9915)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
  PUBMED   14671302
REFERENCE   2 (bases 1 to 9915)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Db   1698 AGACCAAAAGGAACTAAGTGTCAGTGTGCAGTGTTCGACGTCTGGCTGTAAGTGATGGGGTGT 1754

RESULT 11
AY407024
LOCUS   AY407024   9691 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION Mus musculus UTRN gene, VIRTUAL TRANSCRIPT, partial sequence.
            genomic survey sequence.
ACCESSION   AY407024
VERSION   AY407024.1 GI:39762995
KEYWORDS   GSS.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 9691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5552), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 9691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..9691
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene   <1..>9691
                     /gene="UTRN"
                     /locus_tag="HCM2752"
ORIGIN

Query Match   18.9%; Score 301.8; DB 29; Length 9691;
Best Local Similarity   61.3%; Pred. No. 1.5e-57;
Matches 503; Conservative   0; Mismatches 317; Indels   1; Gaps   1;

Qy   104 AGTTCATTGATGGAGGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTA 163
              ||||| |||||| ||| | ||||| |||| | |||| ||||||| |||||||
Db   919 AGCACCGTCAGTGAAGTGGACATGGATTGGACAGCTACCAGATAGCGCTAGGAGGAAGTG 978

Qy   164 TTATCGTGGCTCTTCTTCGTCTGCTAGGACACATTGCAAGCACAAGGAGGAGATTCTAATGAT 223
              ||| || || | || || | |||| |||| ||||| |||||||||| || |||||||
Db   979 CTGACGTGGCTGCTGTCCCGCGAGGACACGTTCCAGGACAAGCAAGATGACATTTCTGATGAT 1038

Qy   224 GTGGAAGTGGTGAAAGACCAGTTTCATACTCATGATGAGGGGTACATGATGATTGACAGCC 283
              |||||||||||||| |||||||| |||| ||||||||||| ||| || |||||||||
Db   1039 GTCGAAGAAGTCAAAGACCAGTTTGCTACCCATGAAACTTTATGATGAGGCTGACAGCA 1098

Qy   284 CATCAGGGCCGGGTTGGTATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAA 343
              |||||||| ||| |||| | ||||| |||||| || ||||||||| || |||
Db   1099 CACCAGGAGCAGCGTGGGGTGGAGCGTCCTGCAGGTCCTGCCACCAGCTGACACCAAGGGACT 1158

Qy   344 TTATCAGAGATGAAGAACTGGAAGTACAAGACAGATACCTAAATTCAAGATGG 403
              ||| ||||| ||||||| ||||| | |||| | ||| | | || |||| ||
Db   1159 CTGTCAGAGGAGGAGGAGTTTGAGATCCAGAACAGATCCAGGGTACAGGTGG 1218

Qy   404 GAATGCCTCAGGGTACCTAGCACGCATGATGCATGGAAAAACAAAGCAATTTACACAGATGGAT 463
              |||| || || || | | | ||| ||||| | ||| | || | |||| |||
Db   1219 GAGGCGGCTCCGGGTGGAGAGAGCCATGGAGAGAGCCAGTCCCGGCTGCACGACGCGCTCTGATGGAG 1278

Qy   464 CTCCAGAATCGA-AACTGAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAGAGAACA 522
              |||| ||||||| ||||||||||||||| ||||||| | | ||| |||||||| |
Db   1279 CTGCAGAAGAACAGCTGCAGCAGCTCTCAAGCTGCTGGCCCTGCCCTCACAGGAGAGCGCATT 1338

Qy   523 AGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAGCCAAGTACAA 582
              ||||| |||||||||| | |||| | | | ||| | | || |||| |||||
Db   1339 CAGAAGATGGAGAGCCCCCCGCTGGGTGATGACCTGCCCTCCCTCCCTGCAGAGCTGCTTCAA 1398

Qy   583 CAACATAAGGTGCTTCAAGAGAGATCAGAAGAGATCTAGAACAAGAACAAGCAGGTCAGGTCAATTCTCTCACT 642
              | ||||||| ||| |||||| || ||| ||| ||| |||||||| || | || ||| ||| | || ||
Db   1399 GAACATAAAGTTGCAAAATGACCTTGAGCTGACAGTGAAGCTGAAGGTAAATTCCTTAACT 1458

Qy   643 CACATGGTGGTGTAGTTGAATCTAGTGGAAGTGGGGAGTGCCACCAACTGCTGTCGTTTGGAAGAA 702
              ||||||||||||| ||||| | | |||| | ||||| | | |||| || || | || ||||||
Db   1459 CCATGGTGGTGATTGTGGATGAAAAACAGTGGGGACGTGCCACAGCAGCTGCGTGTTCGTGAAGAT 1518

Qy   703 CAACTTAAGGTATTGGGAGATCGATGGCAAACATCTGTAGATGGCAGACCGCTGG 762
              || ||| | || |||| | || || ||| | | ||| ||||| |||||||
Db   1519 CAGTTACAGAAACTGGGTGAGCGCTGGGCAGCTGTGATGCCGCTGGACTGAAGACGTTGG 1578

Qy   763 GTTCTTTTACAAGACATCCTCTCAAATGGCAAGCGTCTTACTGAAGAACAGTGCCTTTTT 822
              || |||| | ||| || | ||| ||| | || | ||| || ||| | |||||
Db   1579 AACAGGTTGCAAGAAATCCAGTATTCTGTGGCAGGAATTATTGGAAGAGCAGTGTCTGTTG 1638

Qy   823 AGTGCATGGCGTTTCAGAAAAAGAAGAGATGCAGTGAACAAGATTCACCAACTGGCTTTAAA 882
              |||| | | | |||||||||||||| || |||||| || | ||| || || ||| |||
Db   1639 GAGGCTTGGCTCACCAGGAAAAGGGGAGCCTTTGAATAAAGTTCAAACCAGCAACTTTAAA 1698

Qy   883 GATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGT 923
              | || ||||||| ||| |||| |||| |||||| ||||||
Db   1699 GACCAGAGGAACTAAGTGTCAGTGTCCGGCGTCTGGCTGT 1739

RESULT 12
AY407023
LOCUS   AY407023   9096 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION Pan troglodytes UTRN gene, VIRTUAL TRANSCRIPT, partial sequence.
            genomic survey sequence.
ACCESSION   AY407023
VERSION   AY407023.1 GI:39762994
KEYWORDS   GSS.
SOURCE   Pan troglodytes (chimpanzee)
  ORGANISM Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 9096)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5552), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 9096)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..9096
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene   <1..>9096
                     /gene="UTRN"
                     /locus_tag="HCM2752"
ORIGIN

Query Match   18.7%;   Score 299.8; DB 29; Length 9096;
Best Local Similarity 58.4%;   Pred. No. 4.4e-57;
Matches 489; Conservative   0; Mismatches 347;   Indels   1; Gaps   1;

```
   103 CAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGT 162
       ||||||||||||||||||||  |||||||| ||||||||| ||||||||||||||||||||
   918 CAGCACTGTCACTGAGGTTGACATGGATCTAGACAGCTATCAGATTGCGTTGGAGGAAGT 977

   163 ATTATCGTGGCTTCTTTCTTGCTGAGGACACATTGCAGCAAGGAGAGATTTCTAATGA 222
       ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
   978 GCTGACCTGGTTGCTTCTTCTGCCGAGGACACTTTCCAGGAGCAGGATGATATTTCTGATGA 1037

   223 TGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGC 282
       |||||||| |||||  |||||||||||
   1038 TGTTGAAGAGTCAAAGACCAGTTTGCAACCCATGAANNNNNNNNNNNNNNNNNNNNNNNN 1097

   283 CCATCAGGGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAA 342
       |||||          ||||||
   1098 NNNNCAGAGCAGTGTGGGCAGCGTCCTGCAGGCAACCAACTGATAANNNAAGGAAC 1157

   343 ATTATCAGAAGATGAAGAAACTGAAGTACAAGACGATGAATCTCCTAAATTCAAGATG 402
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  1158 TCTGTCAGANGAAGAGGAATTTGAGATTCAGGAACAGATNNNCCTGCTGAATGCTAGATG 1217

   403 GGAATGCCTCAGGGTAGCTAGCATGGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGA 462
       |||||||||||||||||
  1218 GGAGGCTCTTAGGGTGGAGAGTATGGACANNNAGTCCCGGCTGCACGATGTGTGCTGATGGA 1277

   463 TCTCCAGA-TCGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAAAC 521
       ||||||||
  1278 ACTGCAGAGAGAAGCAACTGCAGCGCTCTCCGCCTGGTTAACACTCACAGAGCGAGCCCAT 1337

   522 AAGGAAAAAATGGAGGAAGAAGAGCCTCTGACCCTGATCTTGAAGACCTAAAACGCCAAGTACA 581
       ||||||||
  1338 TCAGAAGATGGAAACTGCCCCCTGGATGATGATGTAAAATCTCTACAAAAGCTGCTAGA 1397

   582 ACAACATAAGGTGCTTCAGGAGATCTAGAACAGAACAAGTCAGGTCAATTCTCTCAC 641
       ||||||||
  1398 AGAACATAAAGTTTGCAAGTGATCTTGAGGTGAACAGGTGAACAGTAAAGTAAATTCACTAAC 1457

   642 TCAATGGTTGGTAGTTGATGAATCTAGTGGAGATCACGGAACTGCTGCTTTTGGAAGA 701
       ||||||||||||
   762 GGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCCTTTT 821

   762 GGTTCTTTTACAAGACATCCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCCTTTT 821
       ||||||||||
  1578 GAATAGGTTACAAGAAAATCAATATATTGTGGCAGGAATTATTGGAAGAGACACAGTTGCTTGTT 1637

   822 TAGTGCATGGTTTCAGAAAAGAAGATGCAGTCAGTGAACAAGAAGATCCAGACAAGCAACTTCAA 881
       ||||||||||
  1638 GAAAGCTTGGTTAACCGAAAAAAGAAGAGAGGCTTTAAATAAAGTCCAGACAACAAGCAACTTCAA 1697

   882 AGATCAAAATGAAAATGTTATCAAGTGTTATCAAGTGTTAAAAGCGGATCT 938
       ||||||||
  1698 AGACCAAAAGGAACTAACAAGAGTGTCAGTGTTCAGTGTTCGACGTGTTCGGCTGTCGGCTGGGGTTGT 1754
```

REFERENCE
AUTHORS

TITLE   Gene Expression Analysis in Response to Lung Toxicants: I.
        Sequencing and Microarray Development

FEATURES   Location/Qualifiers
  source   1..1122
           /organism="Rattus norvegicus"
           /mol_type="mRNA"
           /strain="Sprague-Dawley"
           /db_xref="taxon:10116"
           /clone="Contig2846"
           /sex="male"
           /tissue_type="airway or parenchyma"
           /dev_stage="adult"
           /clone_lib="Rat lung airway and parenchyma cDNA libraries"
           /note="Organ: lung; Vector: pGEM-11Zf(-); Site 1: Eco RI;
           Site_2: Not I; mRNA was isolated from microdissected rat
           lung airways and parenchyma tissues."

ORIGIN

Query Match   16.5%;   Score 264.8; DB 14; Length 1122;
Best Local Similarity 61.7%;   Pred. No. 2.6e-49;
Matches 439; Conservative   0; Mismatches 272;   Indels   1; Gaps   1;

```
Qy   103 CAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGT 162
         |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   285 CAGCACAGTCAGTCACCGAGGTTGACATGGATTTGGACAGCTACCACCAGTAGCACTAGGAGGAAGT 344

Qy   163 ATTATCGTGGCTTCTTTCTTGCTGAGGACACATTGCAGCAAGGAGAGATTTCTAATGA 222
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   345 ACTGACATGGCTGGTGCTGTCTCGGAGGACACGTTCCAGGAGCAAGATGACATTTCTGATGA 404

Qy   223 TGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGC 282
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   405 TGTCGGAGGACGTCAAAGACCAGTTTGCTACCCATGAACATTGGGAAGTGAGCTGATTGGAACAGGAAA 464

Qy   283 CCATCAGGGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAA 342
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   465 GCACCAGCAGCAGCGTGGGGAGTGTGCTGGCAACACTGATGACACAAGGCAC 524

Qy   343 ATTATCAGAAGATGAAGAAACTGAAGTACAAGACGATGAATCTCCTAAATTCTTCTGATGA 402
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   343 ATTATCAGAAGATGAAGAAACTGAAGTACAAGACGATGAATGAATCTCCTAAATTCAAGATG 402

Qy   403 GGAATGCCTCAGGGTAGCTAGCATGGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGA 462
         |||||||||||||||||||
Db   525 TCTGTCGGATGAGGAGGAGGAGTTTGAGATTCAGGAGCAGCAGTTGAACGCCAGATG 584

Qy   463 TCTCCAGAATCGA-AACTGAAATCGTGAAAATGAGTTTTACATAGAGTTTTAATGA 521
         ||||||||
Db   403 GGAATGCCTCAGGGTAGCTAGCATGCATGAAAGCAATTTACATAGAGTTTTAATGA 462

Qy   522 AAGGAAAAATGGAGGAGGAGGAGCCTCCCAGTTGGTGACGACCTACCCTCCCTGCAAAACCTGCTTGA 584
         |||||||||||
Db   705 TCAGAAGATGGAAAGCCTGCCCAGTTGGGTGACGACCTACCCTCCCTGCAAAACCTGCTTGA 764

Qy   582 ACAACATAAGGTGCTTCAGGAGATCTAGAACAGAACAAGTCAGGTCAATTCTCTCAC 641
         ||||||||
Db   582
```

ESULT 13
F109978
DCUS   CF109978   1122 bp   mRNA   linear   EST 23-JUL-2003
EFINITION Shultzomicao03229 Rat lung airway and parenchyma cDNA libraries
       Rattus norvegicus cDNA clone Contig2846 5', mRNA sequence.
CCESSION CF109978
ERSION   CF109978.1 GI:33165370
EYWORDS   EST.
OURCE   Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
Qy    224 GTGGAAGTGGTGAAAGACCAGTTTCATATCTCATGAGGGGTACATGATGATTTGACAGCC 283
          ||||||||||||||| || || |||| ||||| |||| || ||||  ||| |||||||||
Db    146 GTGCAAGAAGTCAAAGACCAGTTGC-TACCCATGAAACTTTATGGATGGAACTGACAGCA 204

Qy    284 CATCAGGGCCGGTTGGTAATATTCTACAATTGGGAAGTAGCTGATTGGAACAGGAAAA 343
          || || || |  ||| ||||||| ||||||||||||| ||||||||||||||||||||
Db    205 CACCAGCAGCGTGGGACGCTCCTGCAGCGTTGCAACCAGCTGATGACACAAGGGACT 264

Qy    344 TTATCAGAAGATGAAGAAACTGAAGTACAGAGACCAGATGAATCTCCTAAATTCAAGATGG 403
          ||||||||||||||||| ||||| |||| ||||||||||||| |||| |||| |||||| |
Db    265 CTGTCGAGGAGGAGCAGCTGAGATCTGAGATCCAGAAACAAAGCAATTTACATAGAGTTTAATGGAT 324

Qy    404 GAATGCCTCAGGTAGCTAGCATGGAAAAACAAACTGACTGCTAACAAAAAACAGAGAAAGAACA 463
          ||||||| |||||   |||| ||||||| ||| ||| || |||||||| ||||||||||||||
Db    325 GAGGCGGCTCCGGGTGGAGAGCCATGGAGAGGCAGTCCCGGCTCCGGCTCCTGATGGAG 384

Qy    464 CTCCAGAATCGA-AACTGAAAGAGTTGAATGACTGCTAACAAAAGCCAAGTACAA 522
          ||||||||||| |||||||||||||||  | ||||| ||||| ||| || || ||
Db    385 CTGCAGAGAAAAACAGTTGCAGCAGCTGCTCAAGCTGGCTGGCCCTGCTGCCCTCCACAGAGACGCCATT 444

Qy    523 AGGAAAATGGGAGAAGAGCCCTCTTGACCTGTCGATCTTGAGGAGTCCAGACCTAAAACGCCAGTACAA 582
          |||||| |||||||||||||||| | ||| || || |||||||  | ||||| ||| |||||||||||
Db    445 CAGAAGAATGGGAGAGCCCGCCCTCCGTGGGTGATGACCTGCCCTCCCTGCAGAAGCTGCTTCAA 504

Qy    583 CAACATAAGGTGCTTCAGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACT 642
          ||||||||| || ||||||||||| ||| || ||||||||||| || |||| |||||
Db    505 GAACATAAAATTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAACT 564

Qy    643 CACATGGTGGTGTAGTTGATTGGATGAATCTAGTGGAGATCCAACTGCTGCTTTGGAAGAA 702
          |||||||||| |||| | | || || |||  |||||||||| ||||  ||||| ||||||
Db    565 CACATGGTGGTGATTGTGGATGAAAACAGTGGGGAGTGCCACAGCTCCTTCTGAGAAGAT 624

Qy    703 CAACTTAAGGTATTGGGAGGATCGATGGCCAAACATCGTAGATGGCAGAACAGAAGACCGCTGG 762
          |||| || || |||||||| |||| ||||||  || ||| |||||| || |||||| ||| ||
Db    625 CAGTTACAGAAACTGGGTGAGCGCTGGCAGCGTGATGCCGTGCTGACTGAAGAACGTGG 684

Qy    763 GTTCTTTTACAAGACATCCTTCTCAAATGGCAAGTCTTACTGAAGAACAGTGCCTTTT 822
          ||||||| |||||||| || ||||||||||   | ||| ||||||| |||| ||| ||
Db    685 AACAGGTGCAAGAAATCAGTAGTTCTGTGGCCAGGAATTATTGGAAGAGCAGCAGTGTC-TGTA 743

Qy    823 AGTGCATGGCTTTCAGAAAAAGAAGATGC 851
          || ||||| |||||||||| ||||||||
Db    744 GGAGGCTGGATCACCGAAAGGGAAGAGAGGC 772
```

```
RESULT 15
AK081426
LOCUS       AK081426      3753 bp      mRNA      linear      HTC 20-SEP-2003
DEFINITION  Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
            library, clone:C130016K19 product:dystrophin related protein 2,
            full insert sequence.
ACCESSION   AK081426
VERSION     AK081426.1   GI:26349154
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279953
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
```

```
Qy    765 AGAACATAAAAAGTTGCAAAGTGACCTCGAAGCTGACGCAGGTGAAGGTGAATTCCTAAC 824
          |||||||||||| |||||||| |||| || |||||| || ||| ||||||||||||||||
Db    642 TCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCAGTCACGCCAACTGCTGCTTTGGAAGA 701

Qy    825 TCATATGGTGGTGATTGTGGATGAAAAACAGTGGGGAGGAGCCCACAGCTGTTTTGGAAGA 884
          |||||| |||| |||||||||||| ||||||||| ||||| |||| ||||||  |||||||
Db    702 ACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCGTAGATGGCGCTGGACTGGAACCGCTG 761

Qy    885 TCAGTTACAGAAACTGGGTGAGCGCTGGCAGCGTGATGCCGCTGGACTGAAGAACGTTG 944
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    762 GGTTCTTTTACAAGACATCCTTCTCAAATGGCAAGTCTTACTGAAGAACAG 813

Qy    945 GAACAGGTTGCAAGAAATCAATATTTTGTGCAGGAATTATTGGAAGAACAG 996
          |||||| | || |||||||   |||||||| |||||||||||||||||||
Db    814 GAACAGGTTGCAAGAAATCAATATTTTGTGCAGGAATTATTGGAAGAACAG 996
```

```
RESULT 14
BI250598
LOCUS       BI250598      772 bp      mRNA      linear      EST 17-JUL-2001
DEFINITION  602936559F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5149431 5',
            mRNA sequence.
ACCESSION   BI250598
VERSION     BI250598.1   GI:14799101
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 772)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11369 row: e column: 16
            High quality sequence start: 3
            High quality sequence stop: 771.
FEATURES             Location/Qualifiers
     source          1..772
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5149431"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam5"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys
                     NIH"
ORIGIN
Query Match         16.0%;   Score 255.4;   DB 12;   Length 772;
Best Local Similarity   61.4%;   Pred. No. 3.3e-47;
Matches  460;   Conservative  0;   Mismatches 286;   Indels   3;   Gaps   3;
```

```
Qy    104 AGTTCATTGATGGAGGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTA 163
          ||||| ||||   |||||  |||| | |||||||| || ||||||||||| ||||||||
Db    163

Qy    26 AGCACCGTCACTGAGTGGACATGGACATCGATTTGGACAGCTACCAGTACCACCGAGCGAGAAGTG 85
         || || ||| ||||| |||   |  |  ||||| |||| |||   |||   || || || |||| |

Qy    164 TTATCGTGGCTTCTTTCTGCTGAGGACACAATTGCAAGCACACAAGGAGAGATTTCTAATGAT 223
          ||||| |||||||| ||||  |||||||| || |||| | |||| |||||||||||| |||||

Db    86 CTGACGTTGGCTGCTGTGCCCGCGGAGGACACAGTTCCAGGAGCAAGATGACATTTCTGATGAT 145
```

```
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
          Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
          Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
          Yamamoto,R., Matsumoto,K., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
          Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
          Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,K., Kawai,J.,
          Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED    11076861
REFERENCE 4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
          FANTOM Consortium.
TITLE     Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
TITLE     Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE 6  (bases 1 to 3753)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
          Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
          Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
          Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
          Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
          Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
          Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
          Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
          Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
          Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
          Muramatsu,M. and Hayashizaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
          Physical and Chemical Research (RIKEN), Laboratory for Genome
          Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
          RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
          Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
          URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
          Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site for further details.
          URL:http://genome.gsc.riken.go.jp/
          URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..3753
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:C130016K19"
                     /db_xref="MGI:2413800"
                     /db_xref="taxon:10090"
                     /clone="C130016K19"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="16 days embryo"
     CDS             181..3054
                     /note="unnamed protein product; dystrophin related protein
                     2 (MGD|MGI:107432, GB|U43520, evidence: BLASTN, 99%,
                     match=807)
                     putative"
                     /codon_start=1
                     /protein_id="BAC38217.1"
                     /db_xref="GI:26349155"
                     /translation="MQPLVMQGCPYTLPRCHEWHAADRFHHSSSLRNTCPQPQVRAAV
```

```
          TIPAPPWDGAGDPCLSPKLLNGTVGATGPLEPSAMNLCWNEIKKKSHNLRARLEAFSD
          LSGKLQLPLREIIDWLSQKDELSAQLPLQGDVALVQQEKTHAAFMEEVKSKGPYIS
          SVLESAQAFLSQHPFBELEESQSESSKDTSPRQRIQNLSRFVWKQATVASELWEKLTAR
          CVDQHRHIEHTLEHLLEIQGAMEELSSTLTQAEBGVRATWEPIGDLFIDSLPEHIQAIK
          LFKEEFSPVKDGVKLVNDLAHQLAISDVHLSMENSRALEQINIRWKQLQVSVABRLKQ
          LQDAHRDFGPGSQHFLSTSVQVPWERAFSPNKVPYYINHQAQTTCWDHPKMTELYQTL
          ADLNNIKFSAYRTAMKLRRVKALRLDLVTLTTALEIFNEHDLQASEHVMDVVEVIHC
          LTALYERLEEERGILVNVPLCVDMSLAWLLNVFDSGRSGKMRALSFKTGIACLCGTEV
          KEKLQYLPSQVANSGSQCDQRHLGALLHEAIQVPRQLGEVAAFGGSNVEPSVRSCPRF
          STGKPVIEASQPLEWVNLEPQSMVMLAVLHRVTIAEQVKHQTKCSICRQCPIKGFRYR
          SLKQFNVDICQTCFLTGRASKGNKLHYPIMEYYTPTTSSENMRDPATTLKNKFRSKQY
          FSKHPQRGYLPVQSVLESDCSETPASSPMLPHADTHSRIEHFASRLAEMESQNCSFFN
          DSLSPDSIDEDQYLLRHSSPITDREPAFGQAPCSMATESKGELEKILAHLEDENRI
          LQGELRLRLKWQHEEAARAPTLVEGSAEATPDHRNEELLAEARILRQHKSRLETRMQIL
          EDHNKQLESQLQRLRELLLQPPSESDNGSAGSSLLASSPRQSEGSHPREKGQTTPDTE
          VADDVGSKSQDVSLCLEDIMEKLRHAFPSVRSSDVTANTLLAS"
     polyA_signal    3731..3736
                     /note="putative"
     polyA_site      3753
                     /note="putative"
ORIGIN
     Query Match     14.8%;  Score 237.4;  DB 11;  Length 3753;
     Best Local Similarity  67.0%;  Pred. No. 5.4e-43;
     Matches  337;  Conservative   0;  Mismatches 166;  Indels   0;  Gaps   0;

Qy   1099 CAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGAGGCCACGGATGAGCTGGACCTCAAGCTG 1158
          ||| ||||| | || |||| || ||| || ||||| || || || |||||| |||| ||||
Db    862 CATACTCTGGAACATCTATTGGAGATCCAAGGGGCAATGGAGGAATTGAGCAGTACTTTG  921

Qy   1159 CGCCAAGCTGAGGTGATCAAGGGATCCTGCAGCCCGTGGGCGATCTCCTCATTGAGTCT 1218
          | ||| |||| || |||||| ||  |||| ||| ||| || ||||||| || || |||
Db    922 ACCCAAGCAGAGGAGGTCCGAGCCACGTGGGAGCCCATAGGAGATCTCTTTATTGATTCC  981

Qy   1219 CTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAG 1278
          ||||| || || |||| ||||  ||||| ||| || |||||| ||| |||| || ||||
Db    982 CTCCCAGACATATCCAAGCATCAAGCATTCAAAGAGAATTCTCTCGTGTGAAAGAT 1041

Qy   1279 AACGTGAGCCACGTCAATGACCTTGCTCCGCCAGTTACCACTTGGCATTCAGCTCTCA 1338
          |||||| |||| ||| || |||  || |||||| || || ||| ||| || || || |
Db   1042 GGGGTGAAGTTAGTGAATGATCTGGCCCACCAGCTTGCCATCTCGATGTGCACTTGTCA 1101

Qy   1339 CCGTATAACCTGCAGCACTCTGGAGAACCAGATCCGGTGAACACCAGATGGAAGCTTCTCAGGTGGC 1398
          || || |||| ||| || || ||||| || || ||  ||||| | ||||| |||| |||
Db   1102 ATGGAGAATTCCAGGGCTCTGGAACAGATCAACATCCGGTGGAAACAGCTCCAGGTGTCA 1161

Qy   1399 GTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGATCTCAG 1458
          |||||| ||| |||||| || |||  ||| || |||||| || ||||  || |  ||||
Db   1162 GTGGCTGAGAGGCTTAAGCAACTCCAGGATGCCCCACGGGACTTTGGGCTGGTCACAG 1221

Qy   1459 CACTTTCTTCCACGTCTGTCCAGGTCCTGGAGGAGAGAGCCATCTGCGACCAAACAAAGTG 1518
          ||||| ||| ||||  || || || ||||| ||| ||| ||||| ||| |||| || ||
Db   1222 CACTTCCTCCCACTTCTGTCCAAGTTCCCTGGAAAAGCATTTTCTCCCAATAAAGTC 1281

Qy   1519 CCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAG 1578
          ||| || ||| || || |||| ||| || || ||| || ||||| ||||| ||| ||||
Db   1282 CCCTACTACATCAACCACCAGCCTCAGACCCACATGCTGGGACCATCCCAATAAAGTC 1341

Qy   1579 CTCTACCAGTCTTAGCTGACCT 1601
          || || |||| | || || |||
Db   1342 TTATACCAAACCCTAGCCGATCT 1364

Search completed: April  5, 2004, 08:15:21
Job.time : 2804.51 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: April 4, 2004, 16:58:13 ; Search time 77.6428 Seconds
                                    (without alignments)
                                    11443.127 Million cell updates/sec

Title: US-09-845-416-14_COPY_900_2500
Perfect score: 1601
Sequence: 1 gagctatgcctacacacagg........taccagtcttagctgacct 1601

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : Issued Patents NA:*
    1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
    2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
    3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
    4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
    5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
    6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	1092.4	68.2	5952	4	US-09-687-875A-1	Sequence 1, Appli
2	1091	68.1	13977	4	US-09-484-970B-60	Sequence 60, Appl
3	854.2	53.4	19307	3	US-08-836-022A-10.	Sequence 10, Appl
4	854.2	53.4	19307	4	US-09-427-048A-10	Sequence 10, Appl
5	332.6	20.8	6045	4	US-09-091-501B-7	Sequence 7, Appli
6	332.6	20.8	10320	4	US-09-091-501B-9	Sequence 9, Appli
7	234.2	14.6	3915	3	US-09-976-594-93	Sequence 93, Appl
8	79.4	5.0	505	4	US-09-091-501B-5	Sequence 5, Appli
9	78.6	4.9	200	4	US-09-091-501B-4	Sequence 4, Appli
10	78.6	4.9	200	4	US-09-091-501B-6	Sequence 6, Appli
11	74.6	4.7	7218	1	US-08-232-463-14	Sequence 14, Appl
12	42.8	2.7	1179	4	US-09-107-532A-1186	Sequence 1186, Ap
13	42.8	2.7	1690	4	US-09-620-312D-69	Sequence 69, Appl
14	42.8	2.7	2574	4	US-09-668-313A-10	Sequence 10, Appl
15	42.8	2.7	7812	3	US-09-369-590-1	Sequence 1, Appli
16	42.4	2.6	2704	3	US-08-857-076-44	Sequence 44, Appl
17	42.4	2.6	3499	3	US-08-857-076-43	Sequence 43, Appl
18	40.6	2.5	505	4	US-09-621-976-15639	Sequence 15639, A
19	39.4	2.5	832	4	US-09-621-976-2813	Sequence 2813, Ap
20	39.4	2.5	2223	1	US-08-257-073-4	Sequence 4, Appli
21	39.4	2.5	193303	4	US-09-497-855A-37	Sequence 37, Appl
22	39.4	2.5	193303	4	US-09-497-855A-44	Sequence 44, Appl
23	38.8	2.4	289	3	US-09-007-005-17	Sequence 17, Appl
24	38.8	2.4	289	4	US-09-244-796-17	Sequence 17, Appl
25	38.8	2.4	1821	4	US-08-477-831C-1	Sequence 1, Appli
26	38.8	2.4	1885	4	US-08-477-831C-9	Sequence 9, Appli
27	38.8	2.4	1896	4	US-08-477-831C-10	Sequence 10, Appli

28	38.8	2.4	1961	4	US-08-477-831C-8		Sequence 8, Appli
29	38.8	2.4	2968	4	US-08-477-831C-13		Sequence 13, Appli
30	38.8	2.4	3044	4	US-08-477-831C-12		Sequence 12, Appl
31	38.4	2.4	7672	4	US-09-220-132-24		Sequence 24, Appl
32	38.2	2.4	1848	4	US-09-134-001C-447		Sequence 447, App
33	37.4	2.3	2082	3	US-08-985-335-4	c	Sequence 4, Appli
34	37.4	2.3	2082	4	US-09-410-372-4	c	Sequence 4, Appli
35	37.2	2.3	2160	4	US-09-092-218-1		Sequence 1, Appli
36	37.2	2.3	2915	4	US-09-336-115C-5	c	Sequence 5, Appli
37	37.2	2.3	3902	4	US-08-961-527-212		Sequence 212, App
38	36.8	2.3	4929	4	US-09-620-312D-674		Sequence 674, App
39	36	2.2	608	3	US-09-388-982-236		Sequence 236, App
40	36	2.2	2763	3	US-08-248-466-2		Sequence 2, Appli
41	35.8	2.2	633	4	US-09-134-001C-578		Sequence 578, App
42	35.6	2.2	2447	4	US-09-014-969-14		Sequence 14, Appl
43	35.6	2.2	4868	1	US-08-139-937-12		Sequence 12, Appl
44	35.6	2.2	4868	5	PCT-US93-11310-12		Sequence 12, Appl
45	35.6	2.2	8257	4	US-09-595-684B-30		Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6554786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PEI
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 1
; SOFTWARE: PatentIn version 3.1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc_feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match   68.2%; Score 1092.4; DB 4; Length 5952;
Best Local Similarity   99.4%; Pred. No. 0;
Matches 1107; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy   1 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCGCCATT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   891 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCGCCATT 950

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   951 TCCTTCACAGCATTGGAAGCTCCTGAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 1010

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAACAGATTATCGTGGCTTCTTTC 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1011 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAACAGATTATCGTGGCTTCTTTC 1070

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGTGAAAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGTGAAAGA 1130

Qy   241 CCAGTTCATACTCATGAGGGGTACATGATTTGACAGCCCATCAGGGGCCGGGGTTGG 300

```
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match          68.1%;  Score 1091;  DB 4;  Length 13977;
Best Local Similarity 99.9%;  Pred. No. 0;
Matches 1102;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  1;
```

```
Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGCACCTCACACGGAGCCATT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGCACCTCACACGGAGCCATT 1158

Qy    61 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGGCAGTTCATTGCAGTGAGAG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGGCAGTGAGAG 1218

Qy   121 TGAAGTAAACTTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGGCTTCTTTC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 TGAAGTAAACTTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTTCATATCATGAGGGTACATGATTGGACAGCCATCAGGGCCGGGTTGG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTTCATATCATGAGGGGTACATGATTGGACAGCCATCAGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTGGAAGCAGTGAAGATGAAGA  360
           |||||||||||||||||||||||||||||||||||||||
Db  1399 TAATATTCTACAATTGGGAAGTGGAAGCAGTGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459 AACTGAAGTACAAGAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC 1518

Qy   421 TAGCCATGCAAGCAATTTACATAGACAGTTTTAATGGATCTCCCAGAATC-GAAACT  479
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  1519 TAGCCATGCAAGCAATTTACATAGACAGTTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   480 GAAAGAGTTGAATGACTGGCTAACAACAGAGAAAACAGAGAACAAGGAGAAAATGGAGGAAGA  539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579 GAAAGAGTTGAATGACTGGCTAACAACAGAGAAAACAGAGAACAAGGAGAAAATGGAGGAAGA 1638

Qy   540 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACATAAGGTGCTTCA  599
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACATAAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACATGGTGGTGGTAGT  659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACATGGTGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGGATCAAACATCTGTAGATGGACAAGACCGCTGGGTTCTTTTTACAAGACAT  719
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1759 TGATGAATCTAGTGGATCAAACATCTGTAGATGGACAAGACCGCTGGGTTCTTTTTACAAGACAT 1818

Qy   720 AGATCGATGGGCAAACATCTGTAGATGGACAAGACCGCTGGGTTCTTTTTACAAGACAT  779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819 AGATCGATGGGCAAACATCTGTAGATGGACAAGACCGCTGGGTTCTTTTTACAAGACAT 1878

Qy   780 CCTTCTCAAATGCCAACGTCTTACTGAAGACAGTGCCTTTTTTAGTGCATGGCTTTCAGA  839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1879 CCTTCTCAAATGCCAACGTCTTACTGAAGACAGTGCCTTTTTTAGTGCATGGCTTTCAGA 1938

Qy   840 AAAAGAAGATGCAGTGAACACAGATTCACACAACTGCTTTAAAGATCAAAATGAAAATGTT  899
```

```
Db  1131 CCAGTTTCATATCATGAGGGGTACATGATTGGACAGCCATCAGGGCCGGGTTGG 1190

Qy   301 TAATATTCTACAATTGGGAAGTGGAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1191 TAATATTCTACAATTGGGAAGTGGAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA 1250

Qy   361 AACTGAAGTACAAGAGACAGATGAATCTCCTAAATTCCTAAATTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1251 AACTGAAGTACAAGAGACAGATGAATCTCCTAAATTCCTAAATTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC 1310

Qy   421 TAGCCATGGAAAAACAAAGCAATTTACATAGACAGTTTTAATGGATCTCCAGAATC-GAAACT  479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  1311 TAGCCATGGAAAAACAAAGCAATTTACATAGACAGTTTTAATGGATCTCCAGAATCAGAAACT 1370

Qy   480 GAAAGAGTTGAATGACTGGCTAACAACAGAGAAGAAGAACAGAGAAGAAAATGGAGGAAGA  539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1371 GAAAGAGTTGAATGACTGGCTAACAACAGAGAAGAAGAACAGAGAAGAAAATGGAGGAAGA 1430

Qy   540 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACATAAGGTGCTTCA  599
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1431 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCAAGTACAACATAAGGTGCTTCA 1490

Qy   600 AGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACATGGTGGTGTAGT  659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1491 AGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACATGGTGGTGTAGT 1550

Qy   660 TGATGAATCTAGTGGAGATCGGCAACGTCAGGGTCAATTCTCTCGCTTTGGAAGGTATTGGG  719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1551 TGATGAATCTAGTGGAGATCGGCAACGTCAGGGTCAATTCTCTCGCTTTGGAAGGTATTGGG 1610

Qy   720 AGATCCGATGGGCAAACATCTGTAGATGGACAGCACAGATGAATCTTACTGAAGACAGACAT  779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1611 AGATCCGATGGGCAAACATCTGTAGATGGACAGCACAGATGAATCTTACTGAAGACAGACAT 1670

Qy   780 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTTTAGTGCATGGCTTTCAGA  839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1671 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTTTAGTGCATGGCTTTCAGA 1730

Qy   840 AAAAGAAGATGCAGTGAACACAACTCACCACAACTGGCTTCACACAACTGAAATGTT  899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1731 AAAAGAAGATGCAGTGAACACAACTCACCACAACTGGCTTCACACAACTGAAATGTT 1790

Qy   900 ATCAAGTCTTCAAAAACTGCCGTTTAAAGGGGATCTAGAAAGAAGAAAAGCAATCCAT  959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1791 ATCAAGTCTTCAAAAACTGCCGTTTAAAGGGGATCTAGAAAGAAGAAAAGCAATCCAT 1850

Qy   960 GGGCAAACTGTATTCCTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1851 GGGCAAACTGTATTCCTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1910

Qy  1020 CCAGAAGACGGAAGCATGGCTGATAACTTTGCCCGGTAATTAGTCCAAAA 1079
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1911 CCAGAAGACGGAAGCATGGCTGATAACTTTGCCCGGTAATTAGTCCAAAA 1970

Qy  1080 ACTTGAAAAGAGTACAGCACAGACCCTTGAAAGA 1113
           ||||||||||||||||||||||||||||||||||
Db  1971 ACTTGAAAAGAGTACAGCACAGACCACAGGAAGAAAACTGAAATA 2004
```

```
RESULT 2
US-09-484-970B-60
; Sequence No. 6426186, Application US/09484970B
; Patent No. 6426186
GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
```

```
Db  13568  GAGTTATGCCTTCACACAGGCTGCTTATGTTGCCACCTCTGATTCCACACAGAGCCCCTA  13509
Qy     61  TCCTTCACAGCATTTGGAAGCTCCGAAGACAAGCTCATTGGACAGTTCATTGGAGAG        120
           ||| |||||||||||||||||||||||   |||||||| ||||| | ||||||| |||
Db  13508  TCCTTCACAGCATTTGGAAGCTCCCAGAGACACAGTCACTTGACAGTTCATTGATGGGAC  13449
Qy    121  TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGCTTCTTTC     180
           ||||||||| |||| | ||  ||||||| ||||||| |||| ||| | ||||||||||
Db  13448  GGAAGTAAATCTGATAGTTACCAAACTGCTTTAGAAGAATTCATGGCTTCTTTC         13389
Qy    181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA     240
           |||| ||||| |||||||| ||||||||||||||| ||||||| |||| | |||||||
Db  13388  TGCCGAGGATACATTGCGAGCACAAGGAGAGATTCAAATGATGTTGAAGAGTGAAAGA    13329
Qy    241  CCAGTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG       300
           ||||| ||| |||||||||||||| |||||||||||| |||||||| ||| ||||||
Db  13328  ACAGTTCATGCTCATGAGGGATTCATGATGATCTGACACCTCATCACCAAGGACTTGTTGG  13269
Qy    301  TAATATTCTACAATTGGAAGTAAGCTGATTGAACACAGAAAATTATCAGAAGATGAAGA    360
           ||| | |||||||||||||||| | |||| || | ||||||||| |||||||||||||
Db  13268  TAATGTTCTACAGTTAGGAGTCAACTAGTTGGAAAGGGAAATTATCAGAAGATGAAGA    13209
Qy    361  AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCCTAAATTCAAGATGGGAATGTCTCAGGGTAGC  420
Db  13208  AGCTGAAGTGCAAGAACAAATGAATCTCTAAATTCAAGATGGGAATGTCTCAGGGTAGC  13149
Qy    421  TAGCATGTGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT  479
           ||||||||||||||||||| |||||||||| ||||| |||| ||||||| | |||||
Db  13148  TAGCATGTGAAAAACAAAGTCTAATGGATTCTAATGGATCTCCAGAATCAGAAATT     13089
Qy    480  GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAACAAGAAATGAGGAAGA         539
           ||| ||||||||||| ||| |||| ||||||||||||||||||||||||||||
Db  13088  AAAAGAACTAGATGACTGGTTAACAAAAACTGAAGAGAACTAAGAAATGAGGAAGA   13029
Qy    540  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCAAGTACACAACAACTAAGGTGCTTCA  599
Db  13028  GCCCTTTGGACCTGATCTTGAAGATCTAAAATGCCAAGTACACAACAACATAAGGTGCTTCA  12969
Qy    600  AGAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCACTCACATGGTGTGGGTAGT     659
Db  12968  AGAGATCTAGAACAAGAACAGAGCAGGGCGTCAACTGGCTCACATGGTAGTAGTGGT    12909
Qy    660  TGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG  719
Db  12908  TGATGAATCTAGTGGAGATCACGGCAACAGCTGCTTGTTGGAAGAACAACTTAAGGTACTGGG  12849
Qy    720  AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTCTCTTTACAAGACAT   779
Db  12848  AGATCGATGGGCAAATATCTGCAGATGACTGAAGACCGCTGAAGCTGTTTACAAGATAT    12789
Qy    780  CCTTCTCAAATGGCAACGTCTTACTGGAGCAACAGTGCCCTTTTAGTTGCATGCATGG  839
Db  12788  TCTTCTAAAATGGGCAACGTGCTTTGAAGAACAACAGTGCCTTTTAGTTGCATGCATGG   12729
Qy    840  AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT   899
Db  12728  AAAAGAAGATGCAATGAACAATGATTCAGACACTACTTTCGGCACTACTTTCAGTGAC     12669
Qy    900  ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAGAGAAAAAGCAATCCAT   959
Db  12668  GTCAAGTCTTCACAAAATATCTACTTTAAAAATAGTACTTAGAAAAGAGAAAGGCCAACCAT  12609
Qy    960  GGGCAAACTGTATTCCACCACCTCAAACAAGATCTCTTTCAACACTGAAGAGTAAGTCAGTGAC  1019
Db  12608  GGAAAAACTAAGTCACTTTCACTCAATCAATCATAGCTTCGGCACTACTTTCAGTGAC     12549
Qy   1020  CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA   1079
Db  12548  TCAAAGATGGAAATCTGGATGATGGAAAATCTTTGGGATTTTTACCAACTTTAACCCAAA  12489
Qy   1080  ACTTGAAAAGAGTTCAGCACACAGA  1102
Db  12488  ACTTGAAAAGAGTTCAGCACACAAA  12466
```

```
S-08-836-022A-10/c
Sequence 10, Application US/08836022A
Patent No. 6001557
GENERAL INFORMATION:
   APPLICANT: Trustees of the University of Pennsylvania
   APPLICANT: Wilson, James M.
   APPLICANT: Fisher, Krishna J.
   APPLICANT: Chen, Shu-Jen
   APPLICANT: Weitzman, Matthew
   TITLE OF INVENTION: Improved Adenovirus Virus and
   NUMBER OF SEQUENCES: 10
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Howson and Howson
      STREET: Spring House Corporate Cntr, P O Box 457
      CITY: Spring House
      STATE: Pennsylvania
      COUNTRY: USA
      ZIP: 19477
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/836,022A
      FILING DATE:
      CLASSIFICATION: 435
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/331,381
      FILING DATE: 28-OCT-1994
   ATTORNEY/AGENT INFORMATION:
      NAME: Bak, Mary E.
      REGISTRATION NUMBER: 31,215
      REFERENCE/DOCKET NUMBER: GNVPN.008PCT
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: 215-540-9200
      TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 19307 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: unknown
   MOLECULE TYPE: cDNA
S-08-836-022A-10

Query Match             53.4%;  Score 854.2;  DB 3;  Length 19307;
Best Local Similarity   86.5%;  Pred No. 1.2e-252;
Matches  954;  Conservative  0;  Mismatches 148;  Indels 148;  Gaps 1;

Qy     1  GAGCTATGCCTACACACAGGCTGCTTATGTGCCACCACCCTGACCCTACACGAGCGCCATT  60
          || |||||||| ||||||||||||||||| |||||||||||||||| |||||||||||||
Db  12466  ...
```

RESULT 4
US-09-427-048A-10/c
; Sequence 10, Application US/0942704SA
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; Wilson, James M.
; Fisher, Krishna J.
; Chen, Shu-Jen
; Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,048A
; FILING DATE: 21-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,022
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVPN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 53.4%; Score 854.2; DB 3; Length 19307;
Best Local Similarity 86.5%; Pred. No. 1.2e-252;
Matches 954; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

Qy   1 GAGGTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
            || |||||| ||||||||||||||||||||||| ||||||||| |||||||||||
Db 13568 GAGTTATGCCTTCCACACAGGCTGCTATGTTGCCACCTCTGATTCCACACAGAGCCCCTA 13509

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGCCAGTTCATTGATGGAGAG 120
            |||||||||||||||||||||||| |||| ||||||||||||||||||||||||||
Db 13508 TCCTTCACAGCATTTGGAAGCTCCCCAGACAAGTCACTTGACAGTTCATTGATGGAGAC 13449

Qy   121 TGAAGTAAACCTGGACCGTTATCAACAAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
            |||| ||| ||||| ||||| ||| |||| |||||| ||||| |||| ||||||||
Db 13448 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAGTACTTTCATGGCTTCTTTC 13389

Qy   181 TGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAGTGGTGAAAGA 240
            ||| || |||||||||||||| |||||||| |||||||||||||||||| |||||||
Db 13388 TGCCGAGGATACACATTGCGAGCACAAGAGAGATTTCAAATGATGTTGAAGAAGTGAAAGA 13329

Qy   241 CCAGTTCATACTTCATGAGGGGTACATCGATGGATTGCACAGCCCATCAGGGCCGGGTTGG 300
            ||||| ||| |||| ||| |||| ||| |||| |||| |||||||| |||| ||||

PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
  LENGTH: 6045
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (11)..(6037)
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: Chimeric
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (724)..(758)
  OTHER INFORMATION: Precise residue is left open
S-09-091-501B-7

Query Match   20.8%;   Score 332.6; DB 4;   Length 6045;
Best Local Similarity 59.0%;   Pred. No. 8.2e-92;
Matches 589;   Conservative   0; Mismatches 409; Indels   1; Gaps   1;

y   104 AGTTCATTGATGGAGAGTGAAGTAAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTA 163
b   914 AGCACCGTCACTGAAGTGGACATGATTTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 973

y   164 TTATCGTGGCCTCTTCTCGCTGAGGAGACACATTGCAAGACAAGGAGAGAATTCTAATGAT 223
b   974 CTGACGTGGCGTGCTGTCCGGCGGAGGACAGTTCCAGGAGCAAGATGACATTTCTGATGAT 1033

y   224 GTGGAAGTGGTGAAAGACCAGTTCATATCTCATGAGGGGTACATGATGATTTGACAGCC 283
b   1034 GTCGAAGAAGTCAAAGAGCAGTTTTGCTACCCATGAAACTTTTATGATGGAGCTGACAGCA 1093

y   284 CATCAGGGCCCGGGTTGGTAATAATTCTACAAATTGGGAAGTAAGCTAGATGAACAGAAAA 343
b   1094 CACCAGAGCAGCGCGTGGGAGGCGTCCTGCAGGCTGGCACCCAGCTGATGACACAGGGACT 1153

y   344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGG 403
b   1154 CTGTCAGAGGAGGAGGAGGAGTTTGAGATCAGATCCAAGGAACAGATGACCTTGAATGCAAGGTGG 1213

y   404 GAATGCCTCAGGGTAGCTAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTTAATGGAT 463
b   1214 GAGGGCCTCCCGGTGAGAGCATGGAGAGCGCAGTTCCCGGCTGCGCACGCCTCTGATGGAG 1273

y   464 CTCCAGAATCGA-AACTGAAAGAGTGAATGACTGGCTAACAAAAACAGAAGAAGAACA 522
b   1274 CTGCAGAAGAAACAGCTGCAGCAGTTCTCAAGCTGGCTGGCCCCTGGCCCTCACAGAAGAGCGCATT 1333

y   523 AGGAAAATGAGGAGAAGAGCCCTTCTTGGACCTGATCTTGAAGACTCAAAACGCCAAGTACAA 582
b   1334 CAGAAGATGGAGAGAGCGCCCTCCGCTGGGTGATGACTGGCTTCAA 1393

y   583 CAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGACAGTCAGGGTCAATTCTCTCACT 642
b   1394 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGAGTGAAGGTAAATTCCTTAACT 1453

y   643 CACATGGTGGGTGGGTAGTTGATGAATCTAGTGAGATCACGCAACTGCTTGCTTTGGAGAA 702
b   1454 CACATGGTGGGTGGTTGGATTGATGTGGGATGAAAACAGTGCCGGGGAGAGTGCCACAGCTCCTGGAAGAT 1513

y   703 CAACTTAAGGTAGTTGGGAGATTCGAATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGG 762
b   1514 CAGTTACAGAAACTGGGTGAGCGCTGAGCGCGTTATGCGCGCTGGACTGAAGAACGTTGG 1573

y   763 GTTCTTTTACAAGACATCCTTCTCAAATGCAAGTCTTACTGAAGAACCAGTGCTTTT 822
b   1574 AACAGGTTGCAAGAAATATTCTGTGGCCAGAAATATTGGAAGACGAGTCTGTTG 1633

Qy   823 AGTGCATGCGCTTCAGAAAAGAAGAGATGCAGTGAACAGATTCACCACAACTGGCTTTAAA 882
Db   1634 GAGGCTTGGCTCACCGAAAAGGAGGAGAGCGCTTTGAATAAAGTTCAAACCAGCAACTTTAAA 1693

Qy   883 GATCAAAATGAATGTTATCAGTCTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAA 942
Db   1694 GACCAGAAGGAACTAAGTGCAGTGTCAGTGTCCGGCGTCGGCTATATTGAGGAAGCATGAA 1753

Qy   943 AAGAAAAAGCAATCCATGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTG 1002
Db   1754 ATGAAGAGGCAGACTCTGATCAACTGAGTGAGATTGGCCAGGATGTGGGCCAATTACTC 1813

Qy   1003 AAGAATAAGTCAGTGACCCAGAACGGAAGCATGGCTGGGATAACTTTGCCCCGTGTTGG 1062
Db   1814 AGTAATCCCAAGGCATCTAAGAACGATGAACAGTGACTCTGAGGAGCTAACACAGAGATGG 1873

Qy   1063 GATAATTAGTCCAAAACTTGAAAAGAGTACAGCACAG 1101
Db   1874 GATTCTCTGGTTCAGAGACTCGGAGACTCTTCTAACCAG 1912

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
  LENGTH: 10320
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (11)..(10312)
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: Full length
  OTHER INFORMATION: utrophin construct
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match   20.8%;   Score 332.6; DB 4;   Length 10320;
Best Local Similarity 59.0%;   Pred. No. 1.2e-91;
Matches 589;   Conservative   0; Mismatches 409; Indels   1; Gaps   1;

Qy   104 AGTTCATTGATGGAGAGTGAAGTAAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTA 163
Db   914 AGCACCGTCACTGAAGTGGACATGATTTGGACAGCTACCAGATAGCGCTAGAGGAAGTG 973

Qy   164 TTATCGTGGCCTCTTCTCGCTGAGGAGACACATTGCAAGACAAGGAGAGAGATTCTAATGAT 223
Db   974 CTGACGTGGCGTGCTGTCCGGCGGAGGACAGTTCCAGGAGCAAGATGACATTTCTGATGAT 1033

Qy   224 GTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGATTTGACAGCC 283

```
;  SOFTWARE: PERL Program
;  SEQ ID NO 93
;  LENGTH: 3915
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc feature
;  OTHER INFORMATION: Incyte ID No. 6673549 290344.1
US-09-976-594-93

Query Match        14.6%;   Score 234.2;  DB 4;  Length 3915;
Best Local Similarity   68.6%;   Pred. No. 1.6e-61;
Matches  323;  Conservative   0;  Mismatches  148;  Indels   0;  Gaps   0;

Qy   1131 GGCCACGGATGAGCTGACCTCAAGTGCGCCAAGTGAGTGATCAAGGGATCGGCA 1190
          ||||| ||  |||| ||  |||| || ||||||| |||| |||||  |||| ||
Db   1022 GGGCATGGGAGGAACTAAAGCACTACTCTGAGCCAAGCCAAGCTGAGGGAGTCCGAGCCACTTGGGA 1081

Qy   1191 GCCCGTGGGCGATCTCCTCATTGACTCTCCAAGATCACTCGAGAAAGTCAAGGCACT 1250
          || | |||| ||||| ||  ||| | || ||||| ||  ||||  ||  ||||| |||
Db   1082 GCCCATTGGGATCTCTTCATTGATTCACTCCCAGAGCACATCCAGGCTATTAAGCTGTT 1141

Qy   1251 TCGAGGAGAAATTGCGCCTCTGAAACGAACGTGAGCCAGTCAATGACCTTGCTCGCCA 1310
          || | | |  |||  ||| ||| | |||| | |||  ||| ||  ||  |||||  ||
Db   1142 CAAAGAAGAATTCTCCCCCATGAAAGATGGAGTAAAGTTGGTGAATGATCTGGCCACCA 1201

Qy   1311 GCTTACCACTTGGGCATTCAGCTCTCACCGTATAACCTCTGGAAGACCTGAA 1370
          || | |  |||  ||||| |  || || ||  ||| |||| |||||  || ||
Db   1202 ACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAATTCCCAGGCCCTGGAACAGATCAA 1261

Qy   1371 CACCAGATGGAAGCTTCTGCAGTGGCCGTGGAGACCGACTGG 1430
          ||||| ||||||||  | ||| ||| | |||  || ||| ||
Db   1262 CGTCCGATGGAAACAACTACAGGCGTCAGTTGATGAGGATGC 1321

Qy   1431 CCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTCTGTCCAGGGTCCCTG 1490
          || ||||||| ||||| |||  |  ||| |||  | ||  || | || | |  |||| |||
Db   1322 CCACCGGGGACTTTGGGCCTGGGTCACAGCACTTTCTCCTCTGTGTCCAGGGTTCCCTG 1381

Qy   1491 GGAGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACCGAGACTCAAACAAC 1550
          || ||| || ||| |  ||  | ||| | || |||| || |||| |  ||| |||  || ||
Db   1382 GGAAAGAGCAAATTCACCCAATAAAAGTTCCCTACTACATCAACCACCAGGCTCAGGACCAC 1441

Qy   1551 TTGCTGGGACCATCCAAAATGACAGAGCCTCCACCAGTCTTTAGCTGACCT 1601
          |||| | ||| || ||||  ||||||| || |||| ||   ||| ||  ||
Db   1442 ATGCTGGGACCATCCCCAAGATGACAGAGTTTATACCAAACCCTAGCTGATCT 1492


RESULT 8
US-09-091-501B-5
;  Sequence 5, Application US/09091501B
;  Patent No. 6518413
;  GENERAL INFORMATION:
;  APPLICANT: Tinsley, Jonathon M
;  APPLICANT: Davies, Kay E
;  TITLE OF INVENTION: Utrophin gene expression
;  FILE REFERENCE: 620-42
;  CURRENT APPLICATION NUMBER: US/09/091,501B
;  CURRENT FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: PCT/GB96/03156
;  PRIOR FILING DATE: 1996-12-19
;  PRIOR APPLICATION NUMBER: GB 9525962.8
;  PRIOR FILING DATE: 1995-12-19
;  PRIOR APPLICATION NUMBER: GB 9615797.9
;  PRIOR FILING DATE: 1996-07-26
;  PRIOR APPLICATION NUMBER: GB 9622174.2
;  PRIOR FILING DATE: 1996-10-24
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 5
;  LENGTH: 200
;  TYPE: DNA
;  ORGANISM: Rattus sp.
US-09-091-501B-5
```

```
Db   1034 GTCGAAGAAGTCAAAGAGCAGTTGCTACCCATGAAACTTTTATGATGGAGCTGACAGCA 1093

Qy    284 CATCAGGGGCCGGGTTGGTAATATTCTACAAGTAAGCTGATTGGAAGACAGGAAAA 343
          || | |||  ||| ||  ||  |  ||| ||| | |||| | |||||  |||||||
Db   1094 CACCAGAGCAGCGCGTGGGGAGCGTGGGGAGCGCCTGCAGCGTGGCCAACCAGCAACCAGGGACT 1153

Qy    344 TTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGG 403
          || ||||| ||| |||||| ||||  | |||| || ||| || ||| | || ||||| ||
Db   1154 CTGTCAGAGGAGGAGGAGTTTGAGATCCAGGAACAGATGACCTTGCTGAATGCAAGGTGG 1213

Qy    404 GAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGAT 463
          | |||| || || | || |  |||| |||  ||| || | || | |  |||| || ||
Db   1214 GAGGCGCTCCGGGTGGGAGCGCATGGGAGCAGGCAGTCCCGGCTGCACGACGCTCTGATGGAG 1273

Qy    464 CTCAGAATCGA-AACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAGAACA 522
          ||||| ||||  |||| ||||| ||  ||||||  |||| || ||| ||||| ||  |||
Db   1274 CTGCAGAGAAACAGCTGCAGCTCTCAGCTGGCTGGCCCTCACAGAAGAGCGCATT 1333

Qy    523 AGGAAAATGGAGGAAGAGCCTCTTGGACGTCGATCTGAAGACCTAAAACGCCAAGTACAA 582
          |||||||| ||||| || |  || || || | || |||| | ||||| | |||| || ||
Db   1334 CAGAAGATGGAGAGCCCCTCGGTGGGTGATGAGGTCCTGGCTCTGGATGACACTGCTTCAA 1393

Qy    583 CAACATAAGGTGCTTCAGAAGATCCTTACTAGCAACAAGCAAGTCAGGGTCAATTCTCTCACT 642
          |||| | |  ||||| ||  |  || | ||||| |||| | |||| |||| |  | | ||||
Db   1394 GAACATAAAGTTGCAAAATGACCTTGAACAGGTGAAGGTAAAATTCCTTAACT 1453

Qy    643 CACATGGTGGTGGTAGTTGATGAATCTAGTGAGAACACGCAACTGCTGTTTGGAAGAA 702
          ||||| |||||| ||  ||||| ||  ||||  ||||  |  || ||| |||| |||||
Db   1454 CACATGGTGTGATTGTGATGAAAAACAGTGGGGAGAGTGCCACAGCTCTTCTGGAAGAT 1513

Qy    703 CAACTTAAGGTATTGGGAGATCGATGCAAACATCTGTAGATGGACAGAGACCCGCTGG 762
          |||| ||| | |  ||||| ||| |  |  || || | |||||| | || || ||| ||
Db   1514 CAGTTACAGAAAACTGGGTGGTTGAGCGCTGACAGCAGCGTGACTGACGAACGTTGG 1573

Qy    763 GTTCTTTTACAAGACATCCTTCACTGCAAGTCTTACTGAAGAACAGTGCCTTTTT 822
          ||| |||| || | ||| |  | |||  || || |||| ||||| |  | ||||| |
Db   1574 AACAGGTTGCAAGAAATCAGTATTCTGTGGCAGGAATTATTGGAAGAGCAGTGTCTGTTG 1633

Qy    823 AGTGCCATGGCTTTCAGAAGAAGAAGATGCAGTGAACAAGATTCACACAGACTGGCTTAAA 882
          ||||| ||||||  |||||| |||||||||  | |  || ||| | |||| ||| || |||
Db   1634 GAGGCTTGCTCACCGGAAGAGGAGCTTTGAATAAAGTTCAAACCAGCAACTTTAAA 1693

Qy    883 GATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTAAAAAACGGATCTAGAA 942
          || ||||| |||  || | |||  |  | |||||| || | ||  | ||||| ||   ||
Db   1694 GACCAGAAGAACTAAGTGTCAGTGTCCGGCGTGCTCGGCTATATTGAAGGAAGACATGGAA 1753

Qy    943 AAGAAAAAGCAATCCATGGCAACTGTATTCACTCACAAGATCTTCTTTCAACACTG 1002
          || |||| ||| ||  | ||| |  |  |  ||| ||  | |||| || |||| ||
Db   1754 ATGAAGAGGCAGACTCTGGATCAACTGAGTGAGATTGGCCAGGATGTGGGCCAATTACTC 1813

Qy   1003 AAGAATAAGTCAGTGACCCAGAGACGGAAGACGATGGGATAAACTTGCCCGGTGTTGG 1062
          || | |  | ||    || ||||| ||| |||| || |  |   || |   || ||||
Db   1814 AGTAATCCCAAGGCATCTAAGGAGCATCAAGGAGAATGAACAGTGACTCTGAGGAGCTAACAGAGATGG 1873

Qy   1063 GATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACAG 1101
          || ||  || ||  |||||||  ||| || ||
Db   1874 GATTCTCTGGTTCAGAGACTCGAAGACTCTTCAACCAG 1912
```

```
RESULT 7
US-09-976-594-93
;  Sequence 93, Application US/09976594
;  Patent No. 6673549
;  GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT: Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
```

```
Query Match        5.0%;   Score 79.4;  DB 4;  Length 200;
Best Local Similarity  64.3%;  Pred. No. 1.4e-14;
Matches 119;  Conservative  0;  Mismatches  66;  Indels  0;  Gaps  0;

y   564 CCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAAGT 623
            ||||||||||||||||||||||||||||||| ||||||| ||||||
b    16 CCTGCAAAACCTGCTTGAGAACATAAAAGTTTGCAAAGTGACCTCGAAGCTGAGCCAGGT 75

y   624 CAGGGTCAATTCTCTCCACTCACATGGTGTGGTAGTTGATGGAATCTAGTGGAGATCACGC 683
            |||||  |||||
b    76 GAAGGTGAATTCCTTAACTCATATGGTGTGATTGTGATGAAAACAGTGGGGAGAGCGC 135

y   684 AACTGCTGCTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAG 743
               ||||| |||||||
b   136 CACAGCTGTTTGGAAGGATCAGTTACAGAAACTGGGTGAGCCGCTGGACCAGCTGTATGCCG 195

y   744 ATGGA 748
            |||||
b   196 CTGGA 200
```

```
Query Match        4.9%;   Score 78.6;  DB 4;  Length 200;
Best Local Similarity  62.4%;  Pred. No. 2.4e-14;
Matches 123;  Conservative  0;  Mismatches  74;  Indels  0;  Gaps  0;

y   552 TGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCAAGAAGATCTAGA 611
            ||||||  ||                   ||||||||
b     4 TGACCTGCCCTCCCTGCAGAAGCTGCTTCAAGAACAACATAAAAGTTTGCAAAGTGATCTTGA 63

y   612 ACAAGAACAAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTGGTAGTTGATGAATCTAG 671
            ||| |||||
b    64 AGCTGAACAGGTGAAAGTAAATTCCTTAACTCACTAACTCACATGGTGTGATGTGATGAAAACAG 123

y   672 TGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 731
            ||| |||||||
b   124 TGGGGAGAGTGCCACAGCTCTTCTGAAGATCAGTTACAGAAACTGGGTGAGCCGCTGGAC 183

y   732 AAACATCTGTAGATGGA 748
            |||||||||
b   184 AGCTGTATGCCGTGGA 200
```

```
Query Match        4.9%;   Score 78.6;  DB 4;  Length 200;
Best Local Similarity  62.4%;  Pred. No. 2.4e-14;
Matches 123;  Conservative  0;  Mismatches  74;  Indels  0;  Gaps  0;

Qy  552 TGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCAAGAAGATCTAGA 611
            ||||||  ||                   ||||||||
Db    4 TGATGTGAAATCTTCACAAAAGTGCTAGAGAACATAAAAGTTTGCAAAGTGATCTTGA 63

Qy  612 ACAAGAACAAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTGGTAGTTGATGAATCTAG 671
            ||| |||||
Db   64 GGCTGAACAGGTGAAAGTAAATTCACTAACTCACATGGTGGTGATGTGATGATGAAAACAG 123

Qy  672 TGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 731
            ||| |||||||
Db  124 TGGTGAGAGCGCTACAGCTCATCCTAGAAGACCAGTTACAGAAACTTGGTGAGCCGCTGGAC 183

Qy  732 AAACATCTGTAGATGGA 748
            |||||||||
Db  184 AGCAGTATGCCGTTGGA 200
```

```
         FILING DATE:
         APPLICATION NUMBER: EP 91 114 300.6
         FILING DATE: 26-AUG-1991
      ATTORNEY/AGENT INFORMATION:
         NAME: BENT, Stephen A.
         REGISTRATION NUMBER: 29,768
         REFERENCE/DOCKET NUMBER: 30472/114 IMMU
      TELECOMMUNICATION INFORMATION:
         TELEPHONE: (703)836-9300
         TELEFAX: (703)683-4109
         TELEX: 899149
   INFORMATION FOR SEQ ID NO: 14:
      SEQUENCE CHARACTERISTICS:
         LENGTH: 7218 base pairs
         TYPE: nucleic acid
         STRANDEDNESS: single
         TOPOLOGY: linear
      IMMEDIATE SOURCE:
         CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match            4.7%;  Score 74.6; DB 1; Length 7218;
Best Local Similarity  5.6%;  Pred. No. 5e-12;
Matches   23; Conservative  236; Mismatches  150; Indels    0; Gaps     0;

Qy    213 TTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGG 272
              ||||||    ||||    ||||    |   ||    |  :::::: :::: :::::
Db   1474 TATCTATGCAAGTAGTAGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRR 1415

Qy    273 ATTTGACAGCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTG 332
              : :::: :::: ::::  : ::  ::   :: :: :: :: :: :: :: :: :: :
Db   1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355

Qy    333 GAACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAA 392
              ::: ::::::: ::::: :::: :::: :::: :::: :: :::: :::: :::: ::
Db   1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295

Qy    393 ATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGGAAAACAAAAGCAATTACATAGAG 452
              : :::: ::: :::::: :::::: :::: :::: :::: :::: :::::: ::::
Db   1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235

Qy    453 TTTTAATGGATCTCCAGAATCGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGA 512
              :::::: :::: :::: :::: :::: :::::: :::: :::: :::: :: :::
Db   1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175

Qy    513 AGAAAGAACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACG 572
              :::::: :: :::::: :::::: :::: :::: :::: :::: :::: :: :: ::
Db   1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

Qy    573 CCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAA 621
              : ::::::|| :: ::  ::: ::::::|| || |||| :|
Db   1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRA 1066
```

```
         OPERATING SYSTEM: <Unknown>
         SOFTWARE: ASCII
      CURRENT APPLICATION DATA:
         APPLICATION NUMBER: US/09/107,532A
         FILING DATE: 30-Jun-1998
      PRIOR APPLICATION DATA:
         APPLICATION NUMBER: 60/085,598
         FILING DATE: 14 May 1998
         APPLICATION NUMBER: 60/051571
         FILING DATE: July 2, 1997
      ATTORNEY/AGENT INFORMATION:
         NAME: Ariniello, Pamela Deneke
         REGISTRATION NUMBER: 40,489
         REFERENCE/DOCKET NUMBER: GTC-012
      TELECOMMUNICATION INFORMATION:
         TELEPHONE: (781)893-5007
         TELEFAX: (781)893-8277
   INFORMATION FOR SEQ ID NO: 1186:
      SEQUENCE CHARACTERISTICS:
         LENGTH: 1179 base pairs
         TYPE: nucleic acid
         STRANDEDNESS: double
         TOPOLOGY: circular
      MOLECULE TYPE: DNA (genomic)
      HYPOTHETICAL: NO
      ANTI-SENSE: NO
      ORIGINAL SOURCE:
         ORGANISM: Enterococcus faecium
      FEATURE:
         NAME/KEY: misc_feature
         LOCATION: (B) LOCATION 1...1179
         LOCATION:  1...1179
      SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186

Query Match            2.7%;  Score 42.8; DB 4; Length 1179;
Best Local Similarity 49.5%;  Pred. No. 0.0096;
Matches  110; Conservative    0; Mismatches  112; Indels    0; Gaps     0;

Qy    453 TTTTAATGGATCTCCAGAATCGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGA 512
              ||||    |||    ||||    |||  ||||||||||  |||||
Db    585 TGTTGATCCAGCAGTTGAGCAAAGAATAAAGATTTGATCAGATCAGACAAAGAAAAATGG 644

Qy    513 AGAAAGAACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACG 572
              ||||   |||| ||||  |||||||||||||||| ||  ||||
Db    645 AGATACGATCGGAGGAATTGTAGAGTGTCGTCGTTGGAGCGGTTCCAGCTGGATTAGGAAG 704

Qy    573 CCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGCAGGGTCAA 632
              ||| |||||||||||||||  |||||||||||| |||||||||| |||||
Db    705 CTACGTACAATGGACACGAAGCTAGATGCCAAATCGCACAAAATCGCCAAAATCGTGGTTAGTATCAA 764

Qy    633 TTTCTCACTCACATGGGTGGTGGTAGTTGATGAATCTAGTGG 674
              ||||      |||||| ||| ||||||||
Db    765 TGCCTTTAAAGGCGTAGAATTTGGGGTCGGATTCACTTCTGG 806
```

APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 69
LENGTH: 1690
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128)..(1522)
S-09-620-312D-69

Query Match   2.7%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.012; Indels 3; Gaps 1;
Matches 156; Conservative 0; Mismatches 167;

Qy 1116 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTGCGCCAAGCTGAGGTGAT 1175
              ||| | ||| ||| || ||| ||| ||| |||| || ||||| ||
Db   55 CAAGGAGTTGCACCAGGTGCGCCACGACCTGGACGACGAGCTGGCATGGGTTCAGGAGCG 114

Qy 1176 CAAGGGA---TCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCT 1232
Db   115 GCTGCCACTGGCCATGCCATGCAGGACACAGGTTGCAGCGGTCCAGCAGCACAT 174

Qy 1233 CGAGAAAGTCAAGGCACTTCGAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGT 1292
Db   175 CAAAAAGAACCAGGGCCTGCGCGGCGGGAGATCCAGCGCATGGGCGGCCGCGCCTGAGGAGGT 234

Qy 1293 CAATGACCTTGCTCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAG 1352
Db   235 GCTGGAGCGCGGCGGCGGGCGCGTGGCGGTGCGCAGCCCGGAGGCAGAGGCAGTGCGCCG 294

Qy 1353 CACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGT 1412
Db   295 GGGCCCTGGAGCCAGCAGTGCAGAGCGCGCTGCAGGAGCGCCTGGCGGGCGGGAGGCTGCGCGGAGCGCCGGCACGGCA 354

Qy 1413 CAGGCAGCTGCATGAAGCCCACAGGG 1438
Db   355 GCAGGTGCTGGACGCGCGCCTTCCAGG 380

ESULT 14
S-09-668-313A-10
Sequence 10, Application US/09668313A
Patent No. 6503756
GENERAL INFORMATION:
APPLICANT: Brett P. Monia
APPLICANT: Susan M. Freier
APPLICANT: Jacqueline Wyatt
TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
FILE REFERENCE: RTS-0127
CURRENT APPLICATION NUMBER: US/09/668,313A
CURRENT FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 247
SEQ ID NO 10
LENGTH: 2574
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (218)...(1891)

US-09-668-313A-10

Query Match   2.7%; Score 42.8; DB 4; Length 2574;
Best Local Similarity 61.8%; Pred. No. 0.016; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 42;

Qy 1489 TGGGAGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACAGAGACTCAAACA 1548
              |||||||| ||| | |||| ||| ||||||| ||||
Db 1733 TGGGAGGAGGCTTACACAGCAGATGGAATCAAGTACTTCATCAACCACCAGTGACACAGACC 1792

Qy 1549 ACTTCGTCTGGGACCATCCAAAATGACAGAGTCTTACCAGTCTTTAGCTGA 1598
              || | |||| | || ||| | | ||||||
Db 1793 ACGTCCTGGATCCACCCGTGATGAGCGCCCTGAGACCTGTCCTCCTGTGCAGA 1842

RESULT 15
US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
US-09-368-590-1

Query Match   2.7%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.036; Indels 3; Gaps 1;
Matches 156; Conservative 0; Mismatches 167;

Qy 1116 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTGCGCCAAGCTGGCCAAGCTGAGGTGAT 1175
              ||| | ||| ||| || ||| ||| ||| |||| || ||||| ||
Db 3639 CAAGGAGTTGCACCAGGTGCGCCACGACGACCTGGACGACGAGCTGGCATGGGTTCAGGAGCG 3698

Qy 1176 CAAGGGA---TCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCT 1232
Db 3699 GCTGCCACTGGCCATGCCATGCAGGACACAGGTTGCAGCGGTCCAGCAGCACAT 3758

Qy 1233 CGAGAAAGTCAAGGCACTTCGAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGT 1292
Db 3759 CAAAAAGAACCAGGGCCTGCGCGGCGGGAGATCCAGCGCATGGGCGGCCGCGCCTGAGGAGGT 3818

Qy 1293 CAATGACCTTGCTCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAG 1352
Db 3819 GCTGGAGCGCGGCGGCGGGCGCGTGGCGGTGCGCAGCCCGGAGGCAGAGGCAGTGCGCCG 3878

Qy 1353 CACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGT 1412
Db 3879 GGGCCCTGGAGCCAGCAGTGCAGAGCGCGCTGCCAGGAGCGCCTGGCGGGCGGGAGGCTGCGCGGAGCGCCGGCACGGCA 3938

Qy 1413 CAGGCAGCTGCATGAAGCCCACAGG 1438
Db 3939 GCAGGTGCTGGACGCGCGCCTTCCAGG 3964

Mon Apr   5 15:45:52 2004

Search completed: April   5, 2004, 08:25:02
Job time : 79.6428 secs

```
              GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

un on:  April 4, 2004, 11:55:32 ; Search time 413.121 Seconds
                                   (without alignments)
                            16463.377 Million cell updates/sec

itle:          US-09-845-416-14_COPY_900_2500
erfect score:  1601
equence:       1 gagctatgcctacacacagg.........taccagtctttagctgacct 1601

coring table:  IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

arched:        3373863 seqs, 2124099041 residues

otal number of hits satisfying chosen parameters:  6747726

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :  N_Geneseq_29Jan04:*
       1:  geneseqn1980s:*
       2:  geneseqn1990s:*
       3:  geneseqn2000s:*
       4:  geneseqn2001as:*
       5:  geneseqn2001bs:*
       6:  geneseqn2002s:*
       7:  geneseqn2003as:*
       8:  geneseqn2003bs:*
       9:  geneseqn2003cs:*
      10:  geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	DB	Length	ID	Description
1	1601	100.0	6	3446	AAD37242	Human dys
2	1590	99.3	6	4414	AAD37260	Adeno-ass
3	1420.8	88.7	6	5417	ABK81997	DNA encod
4	1193	74.5	6	3510	AAD37240	Human dys
5	1193	74.5	6	4476	AAD37259	Adeno-ass
6	1102	68.8	6	4182	AAD37230	Human dys
7	1102	68.8	6	5149	AAD37255	Adeno-ass
8	1092.4	68.2	5	5952	AAD06794	Human bre
9	1091	68.2	6	8689	ABK82000	DNA encod
10	1091	68.1	6	11241	ABK82005	cDNA enco
11	1091	68.1	6	11443	ABK82002	DNA encod
12	1091	68.1	1	12923	AAN90338	Sequence
13	1091	68.1	6	13957	ABK81959	cDNA enco
14	1091	68.1	6	13957	ABT10904	Human bre
15	1091	68.1	6	13957	ABN95786	Gene #228
16	1091	68.1	6	13957	ABE69900	Human dys
17	1091	68.1	6	13977	ABS70403	Human bon
18	1087.8	67.9	6	11058	AAD37229	Human dys
19	1085.8	67.8	6	1991	AAD37231	Human dys
20	932	58.2	3	4402	AAZ48568	A rod sho
21	869.2	54.3	3	4402	AAZ48567	A rod sho
22	863	53.9	6	3858	AAD37237	Human dys
23	863	53.9	6	4825	AAD37257	Adeno-ass
24	863	53.9	6	4848	AAD37263	Adeno-ass
25	863	53.9	6	5060	AAD37264	Adeno-ass
26	854.2	53.4	6	13815	ABK81960	cDNA enco
27	854.2	53.4	2	13815	ABI99799	Mouse isc
28	854.2	53.4	2	19307	AAT27558	Shuttle v
29	854	53.3	6	5339	ABK81998	DNA encod
30	851.2	53.2	6	3531	AAD37238	Human dys
31	851.2	53.2	6	4498	AAD37258	Adeno-ass
32	846.2	52.9	2	13815	AAV18885	Mus muscu
33	843	52.7	1	3275	AAN97129	Partial s
34	766.4	47.9	6	3999	AAD37234	Human dys
35	766.4	47.9	6	4966	AAD37256	Adeno-ass
36	766.4	47.9	6	4990	AAD37262	Adeno-ass
37	765	47.8	6	1667	AAD37235	Human dys
38	732.2	45.7	6	5462	ABK81999	DNA encod
39	706.8	44.1	3	4075	AAZ48569	A rod sho
40	502.6	31.4	6	1821	AAD37241	Human dys
41	502.6	31.4	6	2169	AAD37232	Human dys
42	501	31.3	6	1434	AAD37243	Human dys
43	450	28.1	6	1340	AAD37239	Human dys
44	409.2	25.6	3	3747	AAZ48566	A rod sho
45	349.4	21.8	7	9195	ACD19399	cDNA enco

ALIGNMENTS

```
RESULT 1
AAD37242
ID   AAD37242 standard; DNA; 3446 BP.
XX
AC   AAD37242;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Human dystrophin minigene delta3447.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 53-54; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
```

:C dystrophy (BMD) in a mammalian subject. The present sequence is human
:C dystrophin minigene delta3447 containing nucleotides 1-1992 (N-terminus,
:C hinge H1 and rods R1, R2 and R3), 8749-10227 (rod R24, hinge H4 and CR
:X domain) and 11047-11058 (dystrophin last 3 amino acids)
:Q Sequence 3446 BP; 1074 A; 766 C; 787 G; 819 T; 0 U; 0 Other;

Query Match   100.0%; Score 1601; DB 6; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GAGCTATGCTACACACCAGGCTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT   60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 GAGCTATGCCTACACACCAGGCTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT  959

Qy         61 TCCTTCACAGCAGATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TCCTTCACAGCAGATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1019

Qy        121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGGCTTCTTTC   180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGGCTTCTTTC  1079

Qy        181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy        241 CCAGTTCATCATGATGAGGGTACATGAGTTTGAATTGCACAGCCCATCAGGGCCCGGGTTGG  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 CCAGTTCATCATGATGAGGGTACATGAGTTTGAATTGCACAGCCCATCAGGGCCCGGGTTGG 1199

Qy        301 TAATATTCTACACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAGATGAAGA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TAATATTCTACACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAGATGAAGA 1259

Qy        361 AACTGAAGTACAAGAGCAGCAGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 AACTGAAGTACAAGAGCAGCAGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy        421 TAGCAATGGGAAAAGCAAATTACATAGAGAGTTTAATGGATCTCCAGAATCGAAACTG  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TAGCAATGGGAAAAGCAAATTACATAGAGAGTTTAATGGATCTCCAGAATCGAAACTG 1379

Qy        481 AAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAGAACAAGGAAGAAATGGAGGAAGAG  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 AAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAGAACAAGGAAGAAATGGAGGAAGAG 1439

Qy        541 CCTCTTGACCTGATCCTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAA  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 CCTCTTGACCTGATCCTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAA 1499

Qy        601 GAAGATCTAGAACAAGAACAAGACAAGTCAGGGTCAATTCTCTCACTCACTGGTGTGGTAGTT  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 GAAGATCTAGAACAAGAACAAGACAAGTCAGGGTCAATTCTCTCACTCACTGGTGTGGTAGTT 1559

Qy        661 GATGAATCTAGTGGAGATCACGCAACTGCTTGGGAAGAACTTAAGGTATTGGGA  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 GATGAATCTAGTGGAGATCACGCAACTGCTTGGGAAGAACTTAAGGTATTGGGA 1619

Qy        721 GATCGATGGGCAAACATCTGTAGAATGGACAGAAGAACCGCTGGGTTCTTTTACAGACATC  780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 GATCGATGGGCAAACATCTGTAGAATGGACAGAAGAACCGCTGGGTTCTTTTACAGACATC 1679

Qy        781 CTTCTCAAATGGCAAGTCTTACTGAAGACAGTGCCTTTTTAGTGCATGGCTTTCAGAA  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 CTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCCTTTTTAGTGCATGGCTTTCAGAA 1739

Qy        841 AAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTAAAGATCAAAATGAAATGTTA  900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAAGAAGATGCAGTGAACAAGATTCACACAACCTGGCTTAAAGATCAAAATGAAATGTTA 1799

Qy        901 TCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAAAAGCAATCCATG  960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 TCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCATG 1859

Qy        961 GGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGACC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1860 GGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGACC 1919

Qy       1021 CAGAAGACGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTAGTCCAAAAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CAGAAGACGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTAGTCCAAAAA 1979

Qy       1081 CTTGAAAAGAGTACAGCACAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGGCCACGGAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 CTTGAAAAGAGTACAGCACAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGGCCACGGAT 2039

Qy       1141 GAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2040 GAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGC 2099

Qy       1201 GATCTCCTCATTGACTCTCCAGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2100 GATCTCCTCATTGACTCTCCAGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGAA 2159

Qy       1261 ATTGCGGCCTCTGAAAGAGAACGTGAGCCACGCACGTCAATGACCTTGCTCGGCCAGCTTACCACT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2160 ATTGCGGCCTCTGAAAGAGAACGTGAGCCACGCACGTCAATGACCTTGCTCGGCCAGCTTACCACT 2219

Qy       1321 TTGGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTGTCGTCCAGGGTCCCTGGGAGAGAGCC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 TTGGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTGTCGTCCAGGGTCCCTGGGAGAGAGCC 2279

Qy       1381 AAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCGAGCCCACAGGGAC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2280 AAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCGAGCCCACAGGGAC 2339

Qy       1441 TTTGGTCCAGCATCTTCAGCACTTTCTTCTTCCACGTGTCGTCCAGGGTCCCTGGGAGAGAGCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2340 TTTGGTCCAGCATCTTCAGCACTTTCTTCTTCCACGTGTCGTCCAGGGTCCCTGGGAGAGAGCC 2399

Qy       1501 ATCTCGCCAAACAAAGTGCCCACTACTATATCAACCACGACTCAAACACTTGCTGGAC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2400 ATCTCGCCAAACAAAGTGCCCACTACTATATCAACCACGACTCAAACACTTGCTGGGAC 2459

Qy       1561 CATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 1601
              |||||||||||||||||||||||||||||||||||||||||
Db       2460 CATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 2500
```

RESULT 2
AAD37260
ID AAD37260 standard; DNA; 4414 BP.
XX
AC AAD37260;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3447.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds;
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN .WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0207777P.
XX

(XIAO/) XIAO X.

Xiao X;

WPI; 2002-049342/06.

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

Example 1; Page 65-66; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence

Sequence 4414 BP; 1255 A; 1086 G; 998 T; 0 U; 0 Other;

Query Match   99.3%;   Score 1590;   DB 6;   Length 4414;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches 1601;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTTCGACCCTCACGACGGAGCCCATT     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTTCGACCCTCACGACGGAGCCCATT   1716

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   1776

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGACAAGTATTATCGTGGCTTCTTTC    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGACAAGTATTATCGTGGCTTCTTTC   1836

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGACAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGGACAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA   1896

Qy    241 CCAGTTCATACTCATGAGGGGTACATGATTGACAGCCCATCAGGGGCCGGGTTGG    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTCATACTCATGAGGGGTACATGATTGACAGCCCATCAGGGGCCGGGTTGG   1956

Qy    301 TAATATTCTACAATTGGGAGTAAGCTAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAGTAAGCTAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA   2016

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   2076

Qy    421 TAGCATGGAGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT    479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGGAGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT   2136

Qy    480 GAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAAACAGAGGAAAATGGAGGAGA    539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAAACAGAGGAAAATGGAGGAGA   2196

Qy    540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA    599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA   2256

Qy    600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT    659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT   2316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    660 TGATGAATCTAGTGGAGATCAGTGGAGATCACCGCAACCAACTGCTGCTTGGAAGAACAACTTGAAGAAGAACAACTTAAGTATTGGG    719

Db   2317 TGATGAATCTAGTGGAGATCACCGCAACCAACTGCTGCTTGGAAGAACAACTTAAGTATTGGG   2376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    720 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCCGTGGGTTCTTTTACAAGACAT    779

Db   2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCCGTGGGTTCTTTTACAAGACAT   2436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA    839

Db   2437 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA   2496
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAAATGTT    899

Db   2497 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAAATGTT   2556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    900 ATCAAGTCTTCAAAAACTGCCGTTTAAAAGCGGATCTAGAAAAGAAGAAAAAGCAATCCAT    959

Db   2557 ATCAAGTCTTCAAAAACTGCCGTTTAAAAGCGGATCTAGAAAAGAAGAAAAAGCAATCCAT   2616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    960 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC   1019

Db   2617 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC   2676
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1020 CCAGAAGACCGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA   1079

Db   2677 CCAGAAGACCGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA   2736
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1080 ACTTGAAAAGAGTACAGACACAGACCCTTGAAAGACTTCAAGGAGCCACGGA   1139

Db   2737 ACTTGAAAAGAGTACAGACACAGACCCTTGAAAGACTTCAAGGAGCCACGGA   2796
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1140 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCTCGGG   1199

Db   2797 TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCTCGGG   2856
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1200 CGATCTCCCATTGACTCCTCCAAGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGA   1259

Db   2857 CGATCTCCCATTGACTCCTCCAAGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGA   2916
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1260 AATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGCTGCCAGCTTACCAC   1319

Db   2917 AATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGCTGCCAGCTTACCAC   2976
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1320 TTTGGGCATTCAGCTCTCCACGTATAACCTCTGGAAGACTCTGAACACCAGATG   1379

Db   2977 TTTGGGCATTCAGCTCTCCACGTATAACCTCTGGAAGACTCTGAACACCAGATG   3036
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1380 GAAGCTTCTGCAGGTGGCCGTGGCGAGTCGCAGCGAGCTGCATGAAGCCCACAGGGA   1439

Db   3037 GAAGCTTCTGCAGGTGGCCGTGGCGAGTCGCAGCGAGCTGCATGAAGCCCACAGGGA   3096
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1440 CTTTGGTCCAGCATCCAGCACTTCTTTCCAGTCTCGTCCCAGGGTCCCTGGAGAGAGC   1499

Db   3097 CTTTGGTCCAGCATCCAGCACTTCTTTCCAGTCTCGTCCCAGGGTCCCTGGAGAGAGC   3156
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1500 CATCTCGCCAAACAAAGTGCCCACTACTATATCAACCACCAGAACTTGCTGGA   1559

Db   3157 CATCTCGCCAAACAAAGTGCCCACTACTATATCAACCACCAGAACTTGCTGGA   3216
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1560 CCATCCAAAAATGACAGAGCTCTACCAGTCTTTAGTGACCT   1601

Db   3217 CCATCCAAAAATGACAGAGCTCTACCAGTCTTTAGTGACCT   3258
```

RESULT 3
ABK81997
ABK81997
ID   ABK81997 standard; DNA; 5417 BP.
XX
AC   ABK81997;

XX
DT   13-AUG-2002 (first entry)
XX
TI   DNA encoding mini-dystrophin protein deltaR4-R23.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200229056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 12; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 U; 0 Other;

Query Match   88.7%; Score 1420.8; DB 6; Length 5417;
Best Local Similarity   91.0%; Pred. No. 0;
Matches 1599; Conservative   0; Mismatches   2; Indels 157; Gaps   2;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTTGGAAGCCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATACTACTGATGAGGGTACATGATGAGGGTACATGGCCCATCAGGGCGGGTTGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTCATACTACTGATGAGGGGTACATGATGAGGGTACATGGCCCATCAGGGCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTGGAAGTTGCTGATTGGAACAGGAGAAAATTATCAGAAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTGGAAGTTGCTGATTGGAACAGGAGAAAATTATCAGAAG 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGATCTCCAGAAATC-GAAACT 479
            |||||||||||||||||||||||||||||||||||||||||||| ||||||| ||||||
Db   1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAAACT 1578

Qy   480 GAAAGAGTTGAATGACTGGCTAACACAAAACAGAGAAAGAACAAGGAAATGAGGAGGAAGA 539
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACACAAAACAGAGAAAGAACAAGGAAATGAGGAGGAAGA 1638

Qy   540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACACATAAGGTGCTTCA 599
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACACATAAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT 659
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGAGAACACGCAACTGTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGGAGATCACGCAACTGTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   720 AGATCGATGGGCAAAATCTCTGTAGATGGACAGAAGAGACCGCTGGGTCTTTTACAGACAT 779
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAAATCTCTGTAGATGGACAGAAGAGACCGCTGGGTTCTTTTACAGACAT 1878

Qy   780 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCTTTTAGTGCATGGCTTCAGA 839
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTTACTGAAGACAGTGCTTTTAGTGCATGGCTTCAGA 1938

Qy   840 AAAAGAAGATGCAGTGAACAGATTCACACAACTGCTTAAAGATCAAATGAAATGTT 899
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAGATTCACACAACTGGCTTAAAGATCAAATGAAATGTT 1998

Qy   900 ATCAAGTCTTCAAAAACTGGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058

Qy   960 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy   1020 CCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
            |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2119 CCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178

Qy   1080 ACTTGAAAAGAGTACAGCACAGA--------------------------- 1102
            ||||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGATTTCACAGGGTGCACCACCACTCAGCCATCACTAAC 2238

Qy   1103 -- 1103

Db   2239 ACAGACAACTGTAATGAAACAGTAACTACGGTGACCACAAGGGAACAGATCCTGGTAAA 2298

Qy   1103 --C 1103

Db   2299 GCATGCTCAAGAGGAACTTCCACCACCTCCCAAAAGAGAGGCAGATTACTGTGGA 2358

Qy   1104 CCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA 1163
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2359 TCTTGAAAGACTTCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCA 2418

Qy   1164 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCCATTGACTCCTCCCA 1223
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2419 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCCATTGACTCCTCCCA 2478

Qy   1224 AGATCACCCTGAAAGTCAAGCACTTCAAGGCCACTTGCGCCTTCGAAAGAATGAAGA 1283
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2479 AGATCACCCTGAAAGTCAAGCACTTCAAGGCCACTTGCGCCTTCGAAAGAATGAAGA 2538

Qy   1284 GAGCCACGTCAATGACCTTGCTCGCCAGCTTCGGCATTCAGCTCTCACCGTA 1343
            ||||||||||||||||||||||||||||||||||||||||||||||||||||

```
2539 GAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTA 2598
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1344 TAACCTCAGCACTCTGGAAGACCTGAACACCAGATGAAGCTTCTGCAGGTGGCCGTCGA 1403
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2599 TAACCTCAGCACTCTGGAAGACCTGAACACCAGATGAAGCTTCTGCAGGTGGCCGTCGA 2658

1404 GGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 1463
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2659 GGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 2718

1464 TCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTA 1523
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2719 TCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTA 2778

1524 CTATATCAACCACGAGACTCAAACAACTGCTGGGGACCATCCCAAAATGACGAGCTCTA 1583
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2779 CTATATCAACCACGAGACTCAAACAACTGCTGGGGACCATCCCAAAATGACGAGCTCTA 2838

1584 CCAGTCTTTAGCTGACCT 1601
     ||||||||||||||||||
2839 CCAGTCTTTAGCTGACCT 2856
```

RESULT 4
AD37240
AD37240 standard; DNA; 3510 BP.

AAD37240;

21-AUG-2002 (first entry)

Human dystrophin minigene delta3510.

Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin; adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD; Becker muscular dystrophy; ds.

Homo sapiens.

WO200183695-A2.

08-NOV-2001.

27-APR-2001; 2001WO-US013677.

28-APR-2000; 2000US-0200777P.

(XIAO/) XIAO X.

Xiao X;

WPI; 2002-049342/06.

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

Example 1; Page 51-52; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus, hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)

```
XX
SQ   Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;

Query Match             74.5%;  Score 1193;  DB 6;  Length 3510;
Best Local Similarity   84.7%;  Pred. No. 0;
Matches 1411;  Conservative   0;  Mismatches  190;  Indels  64;  Gaps   4;

Qy     1 GAGCTATGCTACACACAGGCGTGCTTATGTCACCACCTCTGACCCTACACCGAGCCCATT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCTACACACAGGCGTGCTTATGTCACCACCTCTGACCCTACACCGAGCCCATT 959

Qy    61 TCCTTCACGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAG 120
         |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAG 1019

Qy   121 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAGAAGAGATTTCAAGATGTGGAAGTGGTGAAAGA 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAGAAGAGATTTCAAGATGTGGAAGTGGTGAAAGA 1139

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGCACAAGGAGAGATTTCACAGCCCATCAGGGGCCGGGTTGG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1080 TGCTGAGGACACATTGCAAGCACAAGGAGCACAAGGAGAGATTTCACAGCCCATCAGGGGCCGGGTTGG 1199

Qy   241 CCAGTTCATATCTCATGAGGGGTACATGATGAGTTTGATTTGACAGCCATCAGGGGCCGGGTTGG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1140 CCAGTTCATATCTCATGAGGGGTACATGATGAGTTTGATTTGACAGCCATCAGGGGCCGGGTTGG 1199

Qy   301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200 TAATATTCACAATTGGGAAGTAAGCTGATTGGAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA 1259

Qy   361 AACTGAGTACTACAAGAGCAGCAGATGAATCCTAAATTCAAGATGCCTCAGGGTAGC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260 AACTGAGTACTACAAGAGCAGCAGATGAATCCTAAATTCAAGATGCCTCAGGGTAGC 1319

Qy   421 TAGCCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
         ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  1320 TAGCCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAAACT 1379

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAAATGGGAGGAAGA 539
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1380 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAAATGGGAGGAAGA 1439

Qy   540 GCCTCTTGACCTGATCTTGAGACTTAAACGCCAAGTACAACATAAGGTGCTTCA 599
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1440 GCCTCTTGACCTGATCTTGAGACTTAAACGCCAAGTACAACATAAGGTGCTTCA 1499

Qy   600 AGAAGATCTAGAACAAGACAAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGTGGTAGT 659
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1500 AGAAGATCTAGAACAAGACAAGAACAAGAACAAGTCAGGGTCAATTCTCTCCACTCACATGGTGTGGTAGT 1559

Qy   660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAACAACTTAAGGTATTGGG 719
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACTTAAGGTATTGGG 1619

Qy   720 AGATCGATGGGCAAACATCTGTAGATGGACAAGACCGCTGGGTTCTTTACAAGACAT 779
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1620 AGATCGATGGGCAAACATCTGTAGATGGACAAGACCGCTGGGTTCTTTACAACACAG 1679

Qy   780 CCTTCTCAAATGGCAACGTCTTCTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1680 TTCTGACCAGTGGAGAAGGCTGGAATCAGATCATCTGCGCACCTTTCCTGCAGGGAACTTCTGTGGTGGCTACAGCT 1739

Qy   840 AAAGAAGATGCAGTGAACAAGATTCACAACAGCTGGCTTAAAGATCAAAATGAAATGTT 899
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1740 GAAAGATGATGAATTAAGCGGAGGCACCATTGGAGGCGCACCTATTGGAGGCGCACCTTTTCCAGCGAA 1799

Qy   900 ATCAAGTCTTCAAAACTGGCCGTTTAAAAGGGGATCTAGAAAAGAAAAGCAATCCAT 959
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1800 GCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAAACCTGTAATCAT 1859

Qy   960 GGGCAAACTGTA------------TTCACTCAAACAAGATCTTCTTCAACACTGAAG 1005
         ||| ||| |            ||| ||||| |||||||||
```

```
Db   1860 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 1919

Zy   1006 AATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGA----------------- 1043
                                            |||||
Db   1920 ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGGCT 1979

Zy   1044 -----TAACTTTGCCCGTGTTGGGATAATTTAGTCAAAAACTGAA---------- 1086
               |||| |||                    ||||  ||
Db   1980 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGC 2039

Zy   1087 --------------AAGAGTACAGCACAGACCCCTGAAAGACTCCAGGAACTTCAAGAGGCCAC 1136
                             |||||||||||||||||||||||||||||||||||||||
Db   2040 TGACTGGCAGAGAAAAATAGATGAGACCCCTGAAAGACTCCAGGAACTTCAAGAGGCCAC 2099

Zy   1137 GGATGAGCTGGACCTCAAGCTGCCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGT 1196
               |||||||||||||||||||||||||||||||||  ||||||||||||||||
Db   2100 GGATGAGCTGGACCTCAAGCTGCCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGT 2159

Zy   1197 GGGCGATCTCCCATTGAGTACTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 1256
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2160 GGGCGATCTCCCATTGAGTACTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 2219

Zy   1257 AGAAATTCGCGCCCTCGAAAGAGAACGTGAGCCAGTCAATGCACTTGCTCGCCAGCTTAC 1316
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2220 AGAAATTCGCGCCCTCGAAAGAGAACGTGAGCCAGTCAATGCACTTGCTCGCCAGCTTAC 2279

Zy   1317 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCGGAAGACCTGAACACCAG 1376
               |||||||||||||||||||||||||||||||||||||||||||
Db   2280 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCGGAAGACCTGAACACCAG 2339

Zy   1377 ATGGAAGCTTCTGTCAGGTGGCCGTCGAGGCTGGGCCGTCGAGTCAGGCACGTTCATGATGAAGCCCACAG 1436
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2340 ATGGAAGCTTCTGTCAGGTGGCCGTCGAGGCTGGGCCGTCGAGTCAGGCACGTTCATGATGAAGCCCACAG 2399

Zy   1437 GGACTTGGTCGTCCAGCATCTCAGCACTTCTTCCACGTCTGTGCCAGGGTCCCTGGGAGG 1496
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2400 GGACTTTGGTCGTCCAGCATCTCAGCACTTCTTCCAGTCTGTGCCAGGGTCCCTGGGAGG 2459

Zy   1497 AGCCATCTCGCCAAACAAAGTGCCCTACTTATATCAACCACGAGACTCAAACAACTTGCTG 1556
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460 AGCCATCTCGCCAAACAAAGTGCCCTACTTATATCAACCACGAGACTCAAACAACTTGCTG 2519

Zy   1557 GGACCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCT 1601
               ||||||||||||||||||||||||||||||||||||||||||
Db   2520 GGACCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCT 2564
```

```
Query Match             74.5%;  Score 1193;  DB 6;  Length 4476;
Best Local Similarity   84.7%;  Pred. No. 0;
Matches 1411;      Conservative   0;  Mismatches  190;  Indels    64;  Gaps    4;

Qy      1 GAGCTATGCCTACACACCAGGCTGCTTATGTCACCACCTCTGACCCTCTGACCGCGAGCCCATT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1656 GAGCTATGCCTACACACCAGGCTGCTTATGTCACCACCTCTGACCCTCTGACCCACGCGAGCCCATT 1715

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGGCAGTTCATTGATGGAGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1715 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGGCAGTTCATTGATGGAGAG 1775

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1776 TGAAGTAAACCTGGACCGTTATCAAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1835

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1836 TGCTGAGGACACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1895

Qy    241 CCAGTTCATATCTCATGAGGGGTACATGATGAATTGACAGCCCATCAGGGCCCGGGTTCGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1896 CCAGTTCATATCTCATGAGGGGTACATGATGAATTGACAGCCCATCAGGGCCGGGTTGG 1955

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAAGAAGATGAAGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1956 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAAGAAGATGAAGA 2015

Qy    361 AACTGAAGTACAGAGACAGCAGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2016 AACTGAAGTACAGAGACAGCAGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2075

Qy    421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATC-GAAACT 479
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2076 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGATCTCCAGAATCAGAAAACT 2135

Qy    480 GAAAGAGTTGAATGACTGGCTCAACAAAACAGGAGAACAAGGAGAAAATGGAGGAAGA 539
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2136 GAAAGAGTTGAATGACTGGCTCAACAAAAAACAGAGGAGAAAAATGGAGGAAGA 2195

Qy    540 GCCTCTTGGACCTGATCTTGAAGACGTGACCCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 599
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2196 GCCTCTTGGACCTGATCTTGAAGACCTGAAGACGTGACCCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 2255

Qy    600 AGAAGATCTAGAACAGAACAGCAAGTCAGGGTACTCTCCACTCACGTGGTGGTAGT 659
```

```
AAD37230
ID   AAD37230 standard; DNA; 4182 BP.
XX
AC   AAD37230;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta4173.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 43-44; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
CC   hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC   hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;

Query Match       68.8%;  Score 1102; DB 6;  Length 4182;
Best Local Similarity   96.9%;  Pred. No. 6.1e-305;
Matches 1134;  Conservative   0;  Mismatches  35;  Indels   1;  Gaps   1;

Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCCAAGACAAGTCATTCACCACCTCACGGAGCCCATT   60
Db    900 GAGCTATGCCTACACACAGGCTGCTTATGTCCAAGACAAGTCATTCACCACCTCACGGAGCCCATT  959

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTCATTGGCAGTTCATTGATGGAGAG  120
Db    960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTCATTGGCAGTTCATTGATGGAGAG 1019

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGCAGTTATATCGTGGCTTCTTTC  180
Db   1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGCAGTTATATCGTGGCTTCTTTC 1079

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
Db   1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy    241 CCAGTTCATATCTCATGAGGGGTACAATGAGGGTGATTGACAGCCCATCAGGGCCGGGTTGG  300
```

```
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 57-59; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;

Query Match             68.8%;  Score 1102;  DB 6;  Length 5149;
Best Local Similarity   96.9%;  Pred. No. 6.8e-305;
Matches 1134;  Conservative    0;  Mismatches   35;  Indels    1;  Gaps    1;

Qy      1 GAGCTATGCCTACACACGGCTGCTTATGTCACCACCTCTGACCCTACACCGAGCCCATT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCCTACACACGGCTGCTTATGTCACCACCTCTGACCCTACACCGAGCCCATT 1716

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGCAAGCAGTCATTGGCAGTTCATTGATGGAGAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTGGAAGCTCCTGAAGCAAGCAGTCATTGGCAGTTCATTGATGGAGAG 1776

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGAAGTTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGAAGTTATCGTGGCTTCTTTC 1836

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy    241 CCAGTTTCATACTTCATCATGAGGGGGTACATGAGTGGATTGACAGCCCATCAGGGCCGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTTCATACTTCATCATGAGGGGGTACATGAGTGGATTGACAGCCCATCAGGGCCGGTTGG 1956

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2016
```

```
Db   1140 CCAGTTTCATACTTCATGAGGGGTACAGCCCATCAGCAGGCCGGTTGG 1199
          ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1259

Qy    361 AACTGACAGAGACAGAGCAGATGAATGCCTCAAATTCAAGATGGGATGCCTCAGGGTAGC 420
          ||||||  ||||||||||||||||||||    |||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGATGCCTCAGGGTAGC 1319

Qy    421 TAGGCATGAAAAAACAAAGCAATTTACATAGAGTTTAATGATCTTCCAGAATC-GAAACT 479
          |||| ||||||||||||||||||||||||||||||| ||||||||||||||| ||||||
Db   1320 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1379

Qy    480 GAAAGAGTTGAATGACTGACTGGCTAACAAAAACAGAGAAGAACAAGAGAAAATGGAGGAAGA 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTTGAATGACTGACTGGCTAACAAAAACAGAGAAGAACAAGAGAAAATGGAGGAAGA 1439

Qy    540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACCCAAGTACACACAACATAAGGTGCTTCA 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 GCCTCTTGGACCTGATCTTGAAGACCTAAAAACGCCCAAGTACACACAACATAAGGTGCTTCA 1499

Qy    600 AGAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCTCACTCACATGGTGGTGTAGT 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 AGAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCTCACTCACATGGTGGTGTAGT 1559

Qy    660 TGATGAATCTAGTGGAGATCCAGCAACTGCTGCTTGCTTGGAAGACAACTTAAGGTATTGGG 719
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1560 TGATGAATCTAGTGGAGATCCAGCAACTGCTGCTTGCTTGGAAGACAACTTAAGGTATTGGG 1619

Qy    720 AGATCGATGGGCAAACATCTGTGAGATGAGACAGAAGACCGCTGGGTTCTTTTACAGACAT 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1620 AGATCGATGGGCAAACATCTGTAGATGAGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1679

Qy    780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGTGCCTTTTAGTCGATGCATGGTTCAGA 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1680 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGTGCCTTTTAGTCGATGGCTTTTCAGA 1739

Qy    840 AAAGAAGATGCAGTGAACAAGAATTCACACACTGGCTTAAAGATCAAAAATGAAATGTT 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1740 AAAGAAGATGCAGTGAACAAGAATTCACACACTGGCTTAAAGATCAAAAATGAAATGTT 1799

Qy    900 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1800 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGGCGGATCTAGAAAAAGAAAAAAGCAATCCAT 1859

Qy    960 GGGCAAACTGTATTCACTCAAACAAGACTTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1860 GGGCAAACTGTATTCACTCAAACAAGACTTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1919

Qy   1020 CCAGAAGACGGAAGCATGGCTGGATGAATTTGGCCCGGTGTTGGGATAATTTAGTTGGGAATAA 1079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1920 CCAGAAGACGGAAGCATGGCTGGATGAATTTGGCCCGGTGTTGGGATAATTTAGTTGGGAATAA 1979

Qy   1080 ACTTGAAAAGAGTACAGCAGACACAGACAGACAGCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACCGGA 1139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1980 ACTTGAAAAGAGTACAGCAGACACAGACAGCCATAGATTACTGCAAACAGTTCCCCTGGACCTGGA 2039

Qy   1140 TGAGCTGGACCTCAAGCTGCGCCAAGCTGA 1169
          ||||||||||||||||||||||||||||||
Db   2040 AAAGTTTCTTGCCTGGCTGGCTTACAGAAGCTGA 2069
```

```
RESULT 7
AAD37255
ID   AAD37255 standard; DNA; 5149 BP.
XX
AC   AAD37255;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
```

```
FT   misc_recomb     2847..2848
FT                   /*tag= a
FT                   /label= S4_junction_site
FT                   /note= "Dystrophin gene is split at this site and
FT                   attached to an intein sequence. The resulting fragment is
FT                   used to produce plasmid pSD4"
FT   misc_recomb     2952..2953
FT                   /*tag= b
FT                   /label= S3_junction_site
FT                   /note= "Dystrophin gene is split at this site and
FT                   attached to an intein sequence. The resulting fragment is
FT                   used to produce plasmid pSD3"
FT   misc_recomb     3198..3199
FT                   /*tag= c
FT                   /label= S2_junction_site
FT                   /note= "Dystrophin gene is split at this site and
FT                   attached to an intein sequence. The resulting fragment is
FT                   used to produce plasmid pSD2"
FT   misc_recomb     3300..3301
FT                   /*tag= d
FT                   /label= S1_junction_site
FT                   /note= "Dystrophin gene is split at this site and
FT                   attached to an intein sequence. The resulting fragment is
FT                   used to produce plasmid pSD1"
XX
PN   WO200129243-A1.
XX
PD   26-APR-2001.
XX
PF   13-OCT-2000; 2000WO-CA001216.
XX
PR   15-OCT-1999; 99US-0159868P.
XX
PA   (UYDA-) UNIV DALHOUSIE.
PA   (UYPI-) UNIV PITTSBURGH.
XX
PI   Paul XL, Xiao X;
XX
DR   WPI; 2001-367297/38.
XX
PT   Use of spontaneous or automatic protein splicing to join two or more
PT   peptides at junction site involves expressing extein peptides having co-
PT   reacting portions of split intein attached to them, so that peptides
PT   splice.
XX
PS   Example 1; Fig 2; 81pp; English.
XX
CC   The invention relates to a method directed to the use of spontaneous or
CC   automatic protein trans-splicing to join two or more peptides at junction
CC   site. This method involves expressing extein peptides having co-reacting
CC   portions of split intein attached to them, so that peptides will splice
CC   automatically under suitable conditions. The invention also provides
CC   methods for circumventing virion packaging size limitations in
CC   recombinant virus particle, by splitting a coding region for a protein to
CC   be delivered into two or more extein genes, which are packaged in
CC   separate virus particles and are co-delivered in a target cell for the
CC   expression and for subsequent trans-splicing to form the complete
CC   protein. In particular, the method is used for trans-splicing human
CC   dystrophin and in gene therapies of recombinant adeno-associated virus
CC   (AAV) particles that encode trans-spliced dystrophin, for treating
CC   diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular
CC   dystrophy (BMD). The present sequence is human dystrophin gene (Becker
CC   form)
XX
SQ   Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 U; 0 Other;

Query Match             68.2%;  Score 1092.4;  DB 5;  Length 5952;
Best Local Similarity   99.4%;  Pred. No. 4.2e-302;
Matches  1107;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps     1;

Qy          1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGACCCTACACGGAGCCCATT 60
              |||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        891 GAGCTATGCCTACACACAGGCTGCTTATGTCGTCACCACCCTGACCCTACACGGAGCCCATT 950
```

```
361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATCCTCAGGGTAGC 2076

421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGACTTCCAGAATC-GAAACT 479
        ||||||||||||||||||||||||||||||||||| ||||| |||||||| ||||||
2077 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 2136

480 GAAAGAGTTGAATGACTGGCTAACAGAAAAACAGAAGAAGAACAAGGAAAAATGAGGAAGA 539
        |||||||||||||||||||||||||| |||||||||||||||||||||||||||| ||||
2137 GAAAGAGTTGAATGACTGGCTAACAAAAAAACAGAAGAAGAACAAGGAAAAATGAGGAGGAAGA 2196

540 GCCTCTTGGACCTGATCTTGAAGACTCCAAGTACAACAACATAAGCGTGCTTCA 599
        |||||||||||||||||||||||| ||||||||||||||||||||||||||||
2197 GCCTCTTGGACCTGATCTTGAAGACCTCCAAGTACAACAACATAAGCGTGCTTCA 2256

600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGGTAGT 659
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGTGTAGT 2316

660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2317 TGATGAATCTAGTGGAGATCACGCGAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376

720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACACGCTGGGTTCTTTTACAAGACAT 779
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACACGCTGGGTTCTTTTTACAAGACAT 2436

780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 839
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2437 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCCTTTTTAGTGCATGGCTTTCAGA 2496

840 AAAAGAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 899
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2497 AAAAGAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 2556

900 ATCAAGTCTTCAAAAACTGCCCGTTTAAAAGCGGATCTAGAAAGAGAAAAAGCAATCCAT 959
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2557 ATCAAGTCTTCAAAAACTGCCCGTTTAAAAGCGGATCTAGAAAGAGAAAAAGCAATCCAT 2616

960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2617 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2676

1020 CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTTCCAAAA 1079
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2677 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTTCCAAAA 2736

1080 ACTTGAAAGAGTACAGCACAGACCCTTGAAAGACTCCAGGAGACTTCAAGAGGCCACGGA 1139
        |||||||||||||||||||||||||||| | |||||||||||||||||||||||||||
2737 ACTTGAAAAGAGTACAGCACAGACTACTGCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2796

1140 TGAGCTGGACCTCAGCTCGCGCCAAGCTGA 1169
        |||||||||||||| |||||||||||||||
2797 AAAGTTTCTTGCCTGGCTTACAGAAGCTGA 2826
```

RESULT 8
AD06794
AD06794 standard; DNA; 5952 BP.
XX
AC   AD06794;
XX
DT   06-AUG-2001 (first entry)
XX
DE   Human dystrophin gene (Becker form).
XX
KW   Human; dystrophin; extein; intein; trans-splicing; gene therapy;
KW   Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers

RESULT 9
ABK82000
ID   ABK82000 standard; DNA; 8689 BP.
XX
AC   ABK82000;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaH2-R19.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200229056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 15; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 U; 0 Other;

Query Match   68.1%; Score 1091; DB 6; Length 8689;
Best Local Similarity   99.9%; Pred. No. 1.3e-301;
Matches 1102; Mismatches   0; Conservative   0; Indels   1; Gaps   1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTTCATTGATGGGAGAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTTCATTGATGGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGCAGTTGGAAGTGGTGAAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGCAGTTGGAAGTGGTGAAGA 1338

Qy   241 CCAGTTTCATACTCATGAGGGGTACATGAGGGTACATGATGGATTTGACAGCCCATCAGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCATGAGGGGTACATGAGGGTACATGATGGATTTGACAGCCCATCAGG 1398

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTTCATTGATGGGAGAG 120
Db   951 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTTCATTGATGGGAGAG 1010

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
Db   1011 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1070

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGCAGTTGGAAGTGGTGAAGA 240
Db   1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGCAGTTGGAAGTGGTGAAGA 1130

Qy   241 CCAGTTTCATACTCATGAGGGGTACATGAGGGTACATGATGGATTTGACAGCCCATCAGG 300
Db   1131 CCAGTTTCATACTCATGAGGGGTACATGAGGGTACATGATGGATTTGACAGCCCATCAGG 1190

Qy   301 TAATATTCTACAATTGGAAGTAAGTGATTGGGAACAGCAGAAATATTATCAGAAGATGAAGA 360
Db   1191 TAATATTCTACAATTGGAAGTAAGTGATTGGGAACAGCAGAAATATTATCAGAAGATGAAGA 1250

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db   1251 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1310

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGAATCTCCAGAATC-GAAACT 479
Db   1311 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGAATCTCCAGAATCAGAAACT 1370

Qy   480 GAAAGAGTTGAATGACTGGCTACAAAAACAGAGAGAAGACAAGGAAAATGGGAGGAGA 539
Db   1371 GAAAGAGTTGAATGACTGGCTACAAAAACAGAGAGAAGACAAGGAAAATGGGAGGAGA 1430

Qy   540 GCCTCTTGGACCTGATCTTGAAGACCTAGGGGAAGACCCAAGTACAACACATAAGGTGCTTCA 599
Db   1431 GCCTCTTGGACCTGATCTTGAAGACCTAGGGGAAGACCCAAGTACAACACATAAGGTGCTTCA 1490

Qy   600 AGAAGATCTAGAACAAGAGCAAGTCAGGGTCAATTCTCTCACTGTGGTGGTAGT 659
Db   1491 AGAAGATCTAGAACAAGAGCAAGTCAGGGTCAATTCTCTCACTGTGGTGGTAGT 1550

Qy   660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
Db   1551 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1610

Qy   720 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
Db   1611 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1670

Qy   780 CCCTTCTCAAATGCCAAGCTCTTACTGAAGACAGTGCCTTTTAGTGCATGGCTTCAGA 839
Db   1671 CCTTCTCAAATGCCAAGCTCTTACTGAAGACAGTGCCTTTTAGTGCATGGCTTCAGA 1730

Qy   840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 899
Db   1731 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1790

Qy   900 ATCAAGTCTTCAAAAAACTGGGCCGTTTTAAAAAACGGATCTAGAAAAGAAAAAGCAATCCAT 959
Db   1791 ATCAAGTCTTCAAAAAACTGGGCCGTTTTAAAAAACGGATCTAGAAAAGAAAAAGCAATCCAT 1850

Qy   960 GGGCAAACTGTATTCACTCAAACAAGATCTCTTTCAAACACTGAAGAATAAGTCAGTGAC 1019
Db   1851 GGGCAAACTGTATTCACTCAAACAAGATCTCTTTCAAACACTGAAGAATAAGTCAGTGAC 1910

Qy   1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
Db   1911 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1970

Qy   1080 ACTTGAAAGAGTACAGCACACACCCCTTGAAAGA 1113
Db   1971 ACTTGAAAGAGTACAGCACACAGGAAAACTGAAATA 2004

PN   WO200229056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Example 2; Fig 23; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a human dystrophin polynucleotide
CC   sequence used in the creation of the mini-dystrophin peptides of the
CC   invention
XX
SQ   Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 U; 0 Other;

Query Match   68.1%;   Score 1091; DB 6;   Length 11241;
Best Local Similarity 99.9%;   Pred. No. 1.5e-301;
Matches 1102;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGCAGCGCCATT 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGCAGCGCCCATT 1158

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACAATTGCAGCAGCAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACAATTGCAGCAGCAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTTCATACTTCATGAGGGGTACATGATGGTACATGATGTGACAGCCCATCAGGGCGGGTGG 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTTCATGAGGGGTACATGATGGTACATGATGTGACAGCCCATCAGGGCGGGTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAAACT 1578

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAACAAGAAAGAAATGGAGGAGGA 539
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAACAAGAAAGAAATGGAGGAGGA 1638

Qy   540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACACAACACATAAGGTGCTTCA 599
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACACAACACATAAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT 659
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   720 AGATCGATGGGCAAAACATCGTAGATGGACAGAAGACGCCTGGGTTCTTTACAAGACAT 779
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAAACATCGTAGATGGACAGAAGACGCCTGGGTTCTTTACAAGACAT 1878

Qy   780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTCAGA 839
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTCAGA 1938

Qy   840 AAAAGAAGATGCAGTGAACAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 899
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1998

Qy   900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTCAGAAAAGAAAAGAAATCCAT 959
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTCAGAAAAGAAAAGAAATCCAT 2058

Qy   960 GGGCAAACTGTTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 1019
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGATAAGTCAGTGAC 2118

Qy   1020 CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2178

Qy   1080 ACTTGAAAAGAGTACAGCACGACAGA 1102
             ||||||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACGACAGA 2201

RESULT 10
BK82005
DB   ABK82005 standard; DNA; 11241 BP.
XX
CC   ABK82005;
XX
TD   13-AUG-2002 (first entry)
XX
DE   cDNA encoding human dystrophin, full length HDMD.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; human; ds.
XX
OS   Homo sapiens.

PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 17; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 U; 0 Other;

      Query Match   68.1%;   Score 1091; DB 6;   Length 11443;
      Best Local Similarity   99.9%;   Pred. No. 1.5e-301;
      Matches 1102; Conservative   0; Mismatches   0;   Indels   1;   Gaps   1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCACACGGAGCCCATT   60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAG   120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAGCACAAGGAGAGATTCAAACAGCTTTAGAGAGAAG   180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAGCACAAGGAGAGATTCAAACAGCTTTAGAGAGAAG 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGCACAAGGAGAGATTTCTAATGATGTGGAAAGA   240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGCACAAGGAGAGATTTCTAATGATGTGGAAAGA 1338

Qy   241 CCAGTTTCATACTCCATGAGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG   300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCCATGAGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA   360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGACAAGAACAAGTCAATTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGACAAGAACAAGTCAATTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT   479
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAGAGAAGAACAAGGAAAATGGAGGAAGA   539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAGAGAAGAACAAGGAAAATGGAGGAAGA 1638

Qy   540 GCCCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACACACACATAAGGTGCTTCA   599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACACACACATAAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT   659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGGAGATCACGCAACTGCTCCTTGGAAGAACAACTTAAGGTATTGGG   719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAGATCACGCAACTGCTCCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   720 AGATCGATGGCAAACATCTGTAGATGGAGATCACGCAACATCTGTAGATGGACAAGACAT   779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGCAAACATCTGTAGATGGAGATCACGCAACATCTGTAGATGGACAAGACAT 1878

Qy   540 GCCCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACACACACATAAGGTGCTTCA   599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACACACACATAAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT   659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGGAGATCACGCAACTGCTCCTTGGAAGAACAACTTAAGGTATTGGG   719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAGATCACGCAACTGCTCCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   720 AGATCGATGGCAAACATCTGTAGATGGAGATCACGCAACATCTGTAGATGGACAAGACAT   779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGCAAACATCTGTAGATGGAGATCACGCAACATCTGTAGATGGACAAGACAT 1878

Qy   780 CCTTCTCAAATGGCAACGTCTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA   839
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1938

Qy   840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT   899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998

Qy   900 ATCAAGTCTCAAAAACTGGCCGTTTAAAAACTGGATCTAGAAAAGAAAAGAAAAGCAATCCAT   959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTCAAAAACTGGCCGTTTAAAAACTGGATCTAGAAAAGAAAAGAAAAGCAATCCAT 2058

Qy   960 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCCAACACTCAAGTAAGTAGTCAGTGAC 1019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAACAAGATCTTCTTTCCAACACTCAAGTAAGTAGTCAGTGAC 2118

Qy   1020 CCAGAGACGGAAGCAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAGACGGAAGCAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178

Qy   1080 ACTTGAAAAGAGTACAGCACAGA 1102
             |||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGA 2201

RESULT 11
ABK82002
ID   ABK82002 standard; DNA; 11443 BP.
XX
AC   ABK82002;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaR9-R16.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO20229056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin

Query Match   68.1%;   Score 1091; DB 1;   Length 12923;
Best Local Similarity 99.9%;   Pred. No. 1.6e-301;
Matches 1102;   Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy      1 GAGCTATGCTCACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCTCACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT  1158

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG    120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG   1218

Qy    121 TGAAGTAAACCTGACCGTTATCAAACACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGACCGTTATCAAACACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC  1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  1338

Qy    241 CCAGTTCATACTCATGAGGGGTACAGCAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTCATACTCATGAGGGGTACAGCAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  1398

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  1458

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATCCTCAGGGTAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATCCTCAGGGTAGC  1518

Qy    421 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT   479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1519 TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  1578

Qy    480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGAGGAGGAGA   539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGAGGAGGAGA  1638

Qy    540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACCGCCAAGTACAACAACATAAGGTGCTTCA   599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACCGCCAAGTACAACAACATAAGGTGCTTCA  1698

Qy    600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT   659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT  1758

Qy    660 TGATGAATCTAGTGGAAGCTGCTTGAAAGAAGACAACTTAAGGTATTGGG   719
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAAGCTGCTTGAAAGAAGACAACTTAAGGTATTGGG  1818

Qy    720 AGATGCGATGGGCAAACATCGTGAGATGGACAGAGAGACCGCTGGGTTCTTTTACAAGACAT   779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATGCGATGGGCAAACATCGTGAGATGGACAGAGAGACCGCTGGGTTCTTTTACAAGACAT  1878
```

```
Qy    780 CCTTCCAAATGGCAACGTCTTACTGAGAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA   839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCCAAATGGCAACGTCTTACTGAGAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  1938

Qy    840 AAAGAAGATGCAGTGAACAAGACTCACACAACTGGCTTAAAGATCAAAATGAAATGTT   899
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAGAAGATGCAGTGAACAAGACTCACACAACTGGCTTAAAGATCAAAATGAAATGTT  1998

Qy    900 ATCAAGTCTTCAAAAACTGCCGTTTAAAGACGATCAGAAAAGAAAAAGCAATCCAT   959
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGCCGTTTAAAGACGATCAGAAAAGAAAAAGCAATCCAT  2058

Qy    960 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  1019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  2118

Qy   1020 CCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  1079
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  2178

Qy   1080 ACTTGAAAAGAGTACAGCACAGA  1102
          |||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGA  2201
```

RESULT 12
AN90338

AAN90338 standard; cDNA; 12923 BP.
AAN90338;
29-MAR-1992 (first entry)
Sequence of human muscular dystrophy (MD) cDNA.
Dystrophin; muscular dystrophy; probe; antibody; diagnosis; prenatal; heterozygote; gene therapy; genetic screening; foetal screening; ss.
Homo sapiens.

Key   Location/Qualifiers
CDS   209..12923
         /tag= a

WO890286-A.
13-JUL-1989.
16-DEC-1988; 88WO-US004504.
22-DEC-1987; 87US-00136618.
(CHIL-) CHILDRENS MED CENT.
Kunkel LM, Monaco A, Hoffman EP, Koenig M;
WPI; 1989-220587/30.
P-PSDB; AAP90373.
Muscular dystrophy gene - used for prepn. of probes, dystrophic polypeptide and antibodies for diagnosis and therapy of muscular dystrophy.
Disclosure; Fig 5; 68pp; English.
The inventors claim an MD probe comprising a purified ss NA SQ which hybridises to at least a part of the MD gene; pure dystrophin (DS) polypeptide, purified NA encoding DS and antibodies (Ab) to DS. The probes are equal to or greater than 10b of one of 12 cDNA sequences deposited as ATCC 58666-57677. The MD gene is human, or a murine Dmd gene
Sequence 12923 BP; 4296 A; 2613 C; 2963 G; 3044 T; 0 U; 7 Other;

```
Qy   1020  CCAGAAGACGGAAGCATGGCTGGATAATAACTTTGCCCGGTTGTTGGGATAATTTAGTCCAAAA 1079
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119  CCAGAAGACGGAAGCATGGCTGGATAATAACTTTGCCCGGTTGTTGGGATAATTTAGTCCAAAA 2178

Qy   1080  ACTTGAAAAGAGTACAGCACAGA 1102
           |||||||||||||||||||||||
Db   2179  ACTTGAAAAGAGTACAGCACAGA 2201
```

RESULT 13
ABK81959
ID   ABK81959 standard; DNA; 13957 BP.
XX
AC   ABK81959;
XX
DT   13-AUG-2002 (first entry)
XX
DE   cDNA encoding human dystrophin.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO20229056-A2.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Example 2; Fig 1; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a human dystrophin polynucleotide
CC   sequence used in the creation of the mini-dystrophin peptides of the
CC   invention
XX
SQ   Sequence 13957 BP; 4602 A; 2781 C; 3122 G; 3452 T; 0 U; 0 Other;

Query Match   68.1%; Score 1091; DB 6; Length 13957;
Best Local Similarity   99.9%; Pred. No. 1.7e-301;
Matches 1102; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

Qy   1 GAGCTATGCCTACACAGCCTACACAGGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1099 GAGCTATGCCTACACAGCCTACACAGGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAGAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAGAG 1218

Qy   121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1219 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTGCTTCTTTC 1278
```

---

```
Qy   181  TGCTGAGGACACATTGCAAGCACAAGAGAGAGATTCTAATGATGTGGAAGTGTGAAAGA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279  TGCTGAGGACACATTGCAAGCACAAGAGAGAGATTCTAATGATGTGGAAGTGTGAAAGA 1338

Qy   241  CCAGTTTCATACTCATGAGGGTACATGATTGCACAGCCCATCAGGGGCCGGGTTGG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339  CCAGTTTCATACTCATGAGGGTACATGATTGCACAGCCCATCAGGGGCCGGGTTGG 1398

Qy   301  TAATATTCTACATAATGCAATTACATAGAGTTTAATGATCTCCAGAATC-GAAACT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  1399  TAATATTCTACATAATGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1458

Qy   361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATC-GAAACT  479
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  1519  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1578

Qy   480  GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAGAAGAACAAGGAAAATGGAGGAGA  539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579  GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAGAAGAACAAGGAAAATGGAGGAAGA 1638

Qy   540  GCCTCTTGGACCTGATCTTGAGACCTAAAAGCCAAGTACAACACATAAGGTGCTTCA  599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639  GCCTCTTGGACCTGATCTTGAGACCTAAAAGCCAAGTACAACACATAAGGTGCTTCA 1698

Qy   600  AGAAGATCTAGAACAACAAGAACAAGTCAGGTCAATTCTCCACTCACATGGTCGTGGTAGT  659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699  AGAAGATCTAGAACAACAAGAACAAGTCAGGTCAATTCTCCACTCACATGGTCGTGGTAGT 1758

Qy   660  TGATGAAATCTAGTGGAATCTAGTGGAAGACAACTTAAGGTATTGGG  719
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1759  TGATGAAATCTAGTGGAATCTAGTGGAAGACAACTTAAGGTATTGGG 1818

Qy   720  AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGGCTGGGTTCTTTACAAGACAT  779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819  AGATCGATGGGCAAACATCTGTAGATGACAGAAGACCGGCTGGGTTCTTTACAAGACAT 1878

Qy   780  CCTTCTCAAATGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGCTTTCAGA  839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1879  CCTTCTCAAATGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGCTTTCAGA 1938

Qy   840  AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT  899
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1939  AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAAATGTT 1998

Qy   900  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGGAAAAAGCAATTCCAT  959
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1999  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGGAAAAAGCAATTCCAT 2058

Qy   960  GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAGTCAGTGAC 1019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2059  GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAGTCAGTGAC 2118

Qy  1020  CCAGAAGACGGAAGCATGGCTGGATAATAACTTTGCCCGGTTGTTGGGATAATTTAGTCCAAAA 1079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119  CCAGAAGACGGAAGCATGGCTGGATAATAACTTTGCCCGGTTGTTGGGATAATTTAGTCCAAAA 2178

Qy  1080  ACTTGAAAAGAGTACAGCACAGA 1102
          |||||||||||||||||||||||
Db  2179  ACTTGAAAAGAGTACAGCACAGA 2201
```

RESULT 14
ABT10904
ID   ABT10904 standard; cDNA; 13957 BP.
XX
AC   ABT10904;
XX
DT   04-DEC-2002 (first entry)

```
X
XE  Human breast cancer associated coding sequence SEQ ID NO: 1038.
X
W   Human; breast specific gene; breast cancer; differential expression;
W   cytostatic; gene therapy; gene; ss.
X
S   Homo sapiens.
X
N   WO20025971-A2.
X
D   01-AUG-2002.
X
F   25-JAN-2002; 2002WO-US002176.
X
R   25-JAN-2001; 2001US-0263757P.
R   25-APR-2001; 2001US-0286090P.
R   23-MAY-2001; 2001US-0292517P.
X
A   (GENE-) GENE LOGIC INC.
X
I   Orr MS, Nation M, Diggans JC, Zeng W;
X
R   WPI; 2002-674803/72.
X
T   Diagnosing breast cancer in a patient comprises detecting the level of
T   gene expression in cell or tissue samples, where a differential gene
T   expression is indicative of breast cancer.
X
S   Claim 1; SEQ ID NO 1038; 260pp + Sequence Listing; English.
X
C   The present invention relates to methods of diagnosing breast cancer in a
C   patient, which comprise detecting the level of expression in a tissue
C   sample of two or more genes selected from those shown in ABT09867-
C   ABT11112, where a differential expression of the genes indicates breast
C   cancer. The methods are useful in diagnosing, treating, detecting the
C   progression, and in monitoring treatment of breast cancer in patients.
C   The methods are also useful as a screening tool for agents that modulate
C   the onset or progression of breast cancer. The breast cancer genes may be
C   used as diagnostic markers for the prediction or identification of the
C   malignant state of breast tissue, for confirming the type and progression
C   of cancer, and for drug screening and assays. The present sequence is a
C   coding sequence of the invention. Note: The sequence data for this patent
C   did not form part of the printed specification, but was obtained in
C   electronic format directly from WIPO at
C   ftp.wipo.int/pub.published_pct_sequences
X
Q   Sequence 13957 BP; 4602 A; 2781 C; 3122 G; 3452 T; 0 U; 0 Other;

Query Match            68.1%;  Score 1091;  DB 6;  Length 13957;
Best Local Similarity  99.9%;  Pred. No. 1.7e-301;
Matches 1102;  Conservative 0;  Mismatches 0;  Indels 1;  Gaps 1;
```

```
Qy     361 AACTGAAGTACAAGAGACAGATGAATCTCTAAATTCAAGATGGAATGCCTCAGGGTAGC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    1459 AACTGAAGTACAAGAGACAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy     421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAAACT 1578

Qy     480 GAAAAGTTGAATGACTGGCTAACAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA 539
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    1579 GAAAAGTTGAATGACTGGCTAACAAAACAGAGAAAACAAGGAAAATGGAGGAAGA 1638

Qy     540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCAAGTACAACAACATAAGGTGCTTCA 599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCAAGTACAACAACATAAGGTGCTTCA 1698

Qy     600 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCACTCACATGGTGGTGGTAGT 659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1699 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCACTCACATGGTGGTGGTAGT 1758

Qy     660 TGATGAATCTAGTGAGATCACCGCAACTGCTGCTTGAAGAACAACTTAAGGTATTGGG 719
           ||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||
Db    1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGAAGAACAACTTAAGGTATTGGG 1818

Qy     720 AGATCGATGGGCAAACATCATCTGTAGATGGACAGAAGACCCGCTGGGTTCTTTTACAAGACAT 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1819 AGATCGATGGGCAAACATCATCTGTAGATGGACAGAAGACCCGCTGGGTTCTTTTACAAGACAT 1878

Qy     780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy     840 AAAAGAAGATGCAGTGAACAAGATCACACAACTGGCTTAAAGATCAAAATGAAAATGTT 899
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAAATGTT 1998

Qy     900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGATCTAGAAAAGAAAAGCAATCCAT 959
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGATCTAGAAAAGAAAAAAGCAATCCAT 2058

Qy     960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy    1020 CCAGAAGACGGAAGCAAGCATGGCTGGATAACTTGCCCGGTGTTGGGGATAATTAGTCCAAAA 1079
           |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    2119 CCAGAAGACCGGAAGCAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGGATAATTAGTCCAAAA 2178

Qy    1080 ACTTGAAAAGAGTACAGCACAGA 1102
           |||||||||||||||||||||||
Db    2179 ACTTGAAAAGAGTACAGCACAGA 2201
```

```
RESULT 15
ABN95786
ID   ABN95786 standard; DNA; 13957 BP.
XX
AC   ABN95786;
XX
DT   13-AUG-2002 (first entry)
XX
DE   Gene #2284 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO20022910-A2.
```

```
y     1 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACCTCTGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1099 GAGCTATGCCTACACACAGGCTGCTTATGTCCACCACCTCTGACCCTACACGGAGCCCATT 1158

y    61 TCCTTCCACAGCATTGGAAGCTCCTGAAGACAAGAAGTCATTGGCAGTTCATTGATGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1159 TCCTTCCACAGCATTGGAAGCTCCTGAAGACAAGAAGTCATTGGCAGTTCATTGATGGAGAG 1218

y   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGTATTATCGTGGCTTCTTTC 1278

y   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

y   241 CCAGTTTCATCATGAGGGGTACATGGGTACATGATGATTTGACAGCCCATCAGGCCCGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1339 CCAGTTTCATCATGAGGGGTACATGGGTACATGATGATTTGACAGCCCATCAGGCCCGGGTTGG 1398

y   301 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  1399 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAGAAATTATCAGAAGATGAAGA 1458
```

XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US035089.
XX
PR   02-OCT-2000; 2000US-0237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
PT   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer, hepatocellular
PT   carcinoma or metastatic liver tumor in a patient, involves detecting the
PT   level of expression of two or more genes in a liver tissue sample.
XX
PS   Claim 1; SEQ ID NO 2284; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
KX
SQ   Sequence 13957 BP; 4602 A; 2781 C; 3122 G; 3452 T; 0 U; 0 Other;

Query Match   68.1%; Score 1091; DB 6; Length 13957;
Best Local Similarity   99.9%; Pred. No. 1.7e-301;
Matches 1102; Conservative   .0; Mismatches   0; Indels   1; Gaps   1;

Qy   1 GAGCTATGCCTACACACAGGCGTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCGTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTCGGCAGTCATTGATCGGAGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTCGGCAGTCATTGATCGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTTCATACTCATGAGGGGTACATGAGGGTACATGATTTGACAGTTTGACAGCCCGGGTTGG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCATGAGGGGTACATGAGGGTACATGATTTGACAGTTTGACAGCCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGCAGCAGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGCAGCAGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGACAACAAGGAAAATGGGAGGAAGA 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGACAACAAGGAAAATGGGAGGAAGA 1638

Qy   540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 599
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 719
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 779
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   840 AAAAGAGATGCAGTGAACAAGATTCACACAACTGCTTTAAAGATCAAAATGAAATGTT 899
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAGATGCAGTGAACAAGATTCACACAACTGCTTTAAAGATCAAAATGAAATGTT 1998

Qy   900 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAACTGAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAACTGAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058

Qy   960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy   1020 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCCGTGTTGGGATAATTAGTCCAAAA 1079
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCCGTGTTGGGATAATTAGTCCAAAA 2178

Qy   1080 ACTTGAAAAGAGTACAGCACAGA 1102
           |||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGA 2201

Search completed: April 4, 2004, 18:20:56
Job time : 418.121 secs

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model

un on: April 4, 2004, 12:04:07 ; Search time 4173.93 Seconds
                                     (without alignments)
                            16625.161 Million cell updates/sec

itle: US-09-845-416-14_COPY_900_2500
erfect score: 1601
equence: 1 gagctatgcctacacacagg.......taccagtctttagctgacct 1601

coring table: IDENTITY.NUC
               Gapop 10.0, Gapext 1.0

earched: 3470272 seqs, 2167516995 residues

otal number of hits satisfying chosen parameters: 6940544

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

atabase :   GenEmbl.*
      1: gb_ba.*
      2: gb_htg.*
      3: gb_in.*
      4: gb_om.*
      5: gb_ov.*
      6: gb_pat.*
      7: gb_ph.*
      8: gb_pl.*
      9: gb_pr.*
     10: gb_sts.*
     11: gb_gss.*
     12: gb_sy.*
     13: gb_un.*
     14: gb_vi.*
     15: em_ba.*
     16: em_fun.*
     17: em_hum.*
     18: em_in.*
     19: em_mu.*
     20: em_om.*
     21: em_or.*
     22: em_ov.*
     23: em_pat.*
     24: em_ph.*
     25: em_pl.*
     26: em_ro.*
     27: em_sts.*
     28: em_un.*
     29: em_vi.*
     30: em_htg_hum.*
     31: em_htg_inv.*
     32: em_htg_other:*
     33: em_htg_mus.*
     34: em_htg_pln.*
     35: em_htg_rod.*
     36: em_htg_mam.*
     37: em_htg_vrt:*
     38: em_sy.*
     39: em_htgo_hum.*
     40: em_htgo_mus.*
     41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1420.8	88.7	5417	6	AX538619	AX538619 Sequence
2	1092.4	68.2	5952	6	AR304538	AR304538 Sequence
3	1092.4	68.2	5952	6	AX114289	AX114289 Sequence
4	1091	68.1	8689	6	AX538622	AX538622 Sequence
5	1091	68.1	11443	6	AX538624	AX538624 Sequence
6	1091	68.1	12057	6	AX538627	AX538627 Sequence
7	1091	68.1	13957	6	AX409637	AX409637 Sequence
8	1091	68.1	13957	6	AX538581	AX538581 Sequence
9	1091	68.1	13957	9	HUMDYS	M18533 Homo sapien
10	1091	68.1	13977	9	AR220819	AR220819 Sequence
11	1087.8	67.9	12446	9	HSDMDR	X14298 Human mRNA
12	932	58.2	4402	9	E30220	E30220 Shortened d
13	916.6	57.3	13887	4	AF070485	AF070485 Canis fam
14	866.2	54.3	4402	6	E30219	E30219 Shortened d
15	854.2	53.4	3275	10	MUSDYS	M18025 Mouse dystr
16	854.2	53.4	13815	6	AX306153	AX306153 Sequence
17	854.2	53.4	13815	6	AX538582	AX538582 Sequence
18	854.2	53.4	13815	10	MUSDYSA	M68859 Mouse dystr
c 19	854.2	53.4	19307	6	AR093392	AR093392 Sequence
c 20	854.2	53.4	19307	6	AR142592	AR142592 Sequence
21	854	53.3	5339	6	AX538620	AX538620 Sequence
22	732.2	45.7	5462	6	AX538621	AX538621 Sequence
23	706.8	44.1	4075	6	E30221	E30221 Shortened d
24	623.8	39.0	13575	5	GGDYS	X13369 Chicken mRN
25	612.2	38.2	630	9	HSDMDF1	X06179 Human fetal
26	528	33.0	1966	9	HSDMDA1	X06178 Human adult
27	502.6	31.4	1481	9	BC009242	BC009242 Homo sapi
28	409.2	25.6	3747	6	E30218	E30218 Shortened d
29	349.4	21.8	10302	6	AX538583	AX538583 Sequence
30	349.4	21.8	10302	6	HSMUPS	X69086 H.sapiens m
31	340	21.2	10705	10	RNAJ29967	AJ002967 Rattus no
32	334.2	20.9	1898	10	BC024140	BC024140 Mus muscu
33	332.6	20.8	2341	10	BC062163	BC062163 Mus muscu
34	332.6	20.8	2394	10	BC050183	BC050183 Mus muscu
35	332.6	20.8	6045	6	A63605	A63605 Sequence 7
36	332.6	20.8	6045	6	AR281528	AR281528 Sequence
37	332.6	20.8	10320	6	A63607	A63607 Sequence 9
38	332.6	20.8	10320	6	AR281529	AR281529 Sequence
39	327	20.4	327	6	AX538588	AX538588 Sequence
40	326.2	20.4	6059	6	AX107972	AX107972 Sequence
41	326.2	20.4	11096	6	AX538584	AX538584 Sequence
42	326.2	20.4	11096	10	MMY12229	Y12229 M.musculus
43	324	20.2	324	6	AX538613	AX538613 Sequence
44	322	20.1	333	6	AX538590	AX538590 Sequence
45	321	20.0	333	6	AX538589	AX538589 Sequence

ALIGNMENTS

RESULT 1
AX538619

```
EATURES          Location/Qualifiers
  source         1..5417
                 /organism="synthetic construct"
                 /mol_type="unassigned DNA"
                 /db_xref="taxon:32630"
                 /note="Synthetic"
RIGIN

Query Match      88.7%;   Score 1420.8;   DB 6;   Length 5417;
Best Local Similarity   91.0%;   Pred. No. 0;
Matches 1599;   Conservative  0;   Mismatches  2;   Indels 157;   Gaps  2;
```

```
Qy       1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTGACCCTACACGGAGCCCATT      60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTGACCCTACACGGAGCCCATT    1158

Qy      61 TCCTTCACAGATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG     120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1159 TCCTTCACAGATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG    1218

Qy     121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC    180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   1278

Qy     181 TGCTGAGGACACATTGCAAGCACAAGGAGGAGATTTCTAATGATGTGGAAGTGTGAAAGA    240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1279 TGCTGAGGACACATTGCAAGCACAAGGAGGAGATTTCTAATGATGTGGAAGTGTGAAAGA   1338

Qy     241 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGGCCGGGTTGG    300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1339 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGGCCGGGTTGG   1398

Qy     301 TAATATTCTACAAATTGGGAAGTAAGCTTGATTGGAACAGAGAAAAATTATCAGAAGATGAAGA   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1399 TAATATTCTACAAATTGGGAAGTAAGCTTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA   1458

Qy     361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC    420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   1518

Qy     421 TAGCATGGAAAAACAAAGACAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT    479
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    1519 TAGCATGGAAAAACAAAGACAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   1578

Qy     480 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGGAAGAACAAGGAAAATGAGAGGAAGA    539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1579 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGGAAGAACAAGGAAAATGAGAGGAAGA   1638

Qy     540 GCCTCTTGGACCTGATCTTGAGACAACCTAGTACAACGCCAGTACAACAACATAAGGTGCTTCA    599
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1639 GCCTCTTGGACCTGATCTTGAGACAACCTAGTACAACGCCAGTACAACAACATAAGGTGCTTCA   1698

Qy     600 AGAAGATCTAGAACAAGAACAAGATTCACACAACTTCACTCACTTGGTGGTGGTAGT    659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1699 AGAAGATCTAGAACAAGAACAAGATTCACACAACTTCACTCACTTGGTGGTGGTAGT   1758

Qy     660 TGATGAATCTAGTGGAGCAAACATCTGTAGATGGACAGCAGACACCGCTGGGTTCTTTACCAACACAT    719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1759 TGATGAATCTAGTGGAGCAAACATCTGTAGATGGACAGCAGACACCGCTGGGTTCTTTACCAACACAT   1818

Qy     720 AGATCGATGGGCAAACATCTGTAGATGGACAGCAGACACCGCTGGGTTCTTTTAGTGTAGACAT    779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1819 AGATCGATGGGCAAACATCTGTAGATGGACAGCAGACACCGCTGGGTTCTTTTAGTGTAGACAT   1878

Qy     780 CCTTTCCAAATGGCAACGTCTTACTGAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA    839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1879 CCTTTCCAAATGGCAACGTCTTACTGAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA   1938

Qy     840 AAAAGAAGATGCAGTGAACAAGCAAGATTCACACAACTGGCTTTAAAGATCAAAAATGCTT    899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1939 AAAAGAAGATGCAGTGAACAAGCAAGATTCACACAACTGGCTTTAAAGATCAAAAATGTTT   1998

Qy     900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGGATCTAGAAAAGAAAAAGCAATCCAT    959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGGGGATCTAGAAAAGAAAAAGCAATCCAT    2058

Qy     960 GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC    1019
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2059 GGGCAAACTGTATTCACTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC    2118

Qy    1020 CCAGAAGACGGAAGCATGGTGGATAAACTTGCCCGTGTTGGGATAATTTAGTCCAAAA    1079
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2119 CCAGAAGACGGAAGCATGGTGGATAAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA   2178

Qy    1080 ACTTGAAAAGAGTACAGCACAGA----------------------------------------------    1102
           ||||||| ||  ||||||
Db    2179 ACTTGAAAAGAGTACAGCACAGATTTCACAGGCTGTCACCACCACTCAGCCATCACTAAC   2238

Qy    1103 -------------------------------------------------------------    1102
Db    2239 ACAGAACAACTGTAATGGAAACAGTAACTACGGTGACCACAAGGGAACAGATCCTGGTAAA   2298

Qy    1103 ----------------------------------------------------------C   1103
Db    2299 GCATGCTCAAGAGGAACTTCCACCACCACCTCCCAAAAGAGAGGCAGAGATTACTGTGGA   2358

Qy    1104 CCTTGAAAGACTCCAGGAACTTCAAGAGAGCCACGGATGAGCTGGACTTCAAGCTGCGCCA    1163
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2359 TCTTGAAAGACTCCAGGAACTTCAAGAGAGCCACGGATGAGCTGGACTTCAAGCTGCGCCA   2418

Qy    1164 AGCTGAGGTGATCAAGGATCCTGGCAGCCCATCTCCTCATTGACTCTCTCCA    1223
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2419 AGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGTGGGCCATCTCCTCATTGACTCTCTCCA   2478

Qy    1224 AGATCACCTCGAGAAGAGTCAAGGACACTTCGAGACGAAATTGGCGCCTCTGAAAGAGAACGT    1283
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2479 AGATCACCTCGAGAAGAGTCAAGGACACTTCGAGACGAAATTGCGCGCCTCTGAAAGAGAACGT   2538

Qy    1284 GAGCCAGTCGCAATGACCTTGCTCCGCCAGCTTGCTCCGCCAGCTTTGGCATTCAGCACTT    1343
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2539 GAGCCAGTCGCAATGACCTTGCTCCGCCAGCTTGCTCCGCCAGCTTTGGCATTCAGCACTT   2598

Qy    1344 TAACCTGACACTCTGGAAGACTGCATGAGAGCTGCATGAGAGCACTTTCGCGAGGTGGCCGTCGA    1403
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2599 TAACCTGACACTCTGGAAGACTGCATGAGAGCTGCATGAGAGCACTTTCGCGAGGTGGCCGTCGA   2658

Qy    1404 GGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGCACTT    1463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2659 GGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGCACTT   2718

Qy    1464 TCTTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTA    1523
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2719 TCTTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTA   2778

Qy    1524 CTATATCAACCACCAGAGACTCAAACAACTGCTGGGACCATCCCAAAATGACAGAGCTCTA    1583
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2779 CTATATCAACCACCAGAGACTCAAACAACTGCTGGGACCATCCCAAAATGACAGAGCTCTA   2838

Qy    1584 CCAGTCTTTAGCTGACCT    1601
           ||||||||||||||||||
Db    2839 CCAGTCTTTAGCTGACCT    2856
```

```
RESULT 2
AR304538
LOCUS       AR304538        5952 bp       DNA        linear     PAT 12-JUN-2003
DEFINITION  Sequence 1 from patent US 6544786.
ACCESSION   AR304538
VERSION     AR304538.1    GI:31693691
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 5952)
  AUTHORS   Xiao,X. and Liu,P.X.
  TITLE     Method and vector for producing and transferring trans-spliced
```

```
JOURNAL    peptides
FEATURES   Patent: US 6544786-A 1 08-APR-2003;
source     1. .5952
           /organism="unknown"
           /mol_type="genomic DNA"
RIGIN

Query Match          68.2%;     Score 1092.4;  DB 6;   Length 5952;
Best Local Similarity 99.4%;    Pred. No. 3e-257;
Matches 1107;  Conservative   0;  Mismatches   6;  Indels   6;  Gaps   1;

Qy       1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     891 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   950

Qy      61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTCATTGGCAGTCATTGGCAGTCATTGATGGAGAG   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     951 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTCATTGATGGAGAG   1010

Qy     121 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1011 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC   1070

Qy     181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGGTGGTGAAAGA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGGTGGTGAAAGA   1130

Qy     241 CCAGTTTCATACTTCATGAGGGGTACAGCCCATCAGCGGCCGGGTTGG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1131 CCAGTTTCATACTTCATGAGGGGTACAGCCCATCAGCGGCCGGGTTGG   1190

Qy     301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1191 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA   1250

Qy     361 AACTGAAGTACAAGAGCAGAATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1251 AACTGAAGTACAAGAGCAGAATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC   1310

Qy     421 TAGCATGAAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATC-GAAACT   479
           |||||||||||||||||||||||||||||||||||| |||||||||||||| ||||||
Db    1311 TAGCATGAAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT   1370

Qy     480 GAAAGAGTTGAATGACTGGCTAACAATGCTGCTAACAAAAACAGAGAAGAACAAGGAGAAGA   539
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1371 GAAAGAGTTGAATGACTGGCTAACAATGCTGCTAACAAAAACAGAGAAGAACAAGGAGAAGA   1430

Qy     540 GCCTCTTGGACCTGCATCTTGAGACCTCTGAAGAACCCAAGTACAACACCAAGGTGCTTCA   599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1431 GCCTCTTGGACCTGCATCTTGAGACCTCTGAAGAACCCAAGTACAACACCAAGGTGCTTCA   1490

Qy     600 AGAGATCTAGAACAAGAACAAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT   659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1491 AGAGATCTAGAACAAGAACAAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT   1550

Qy     660 TGATGAATCTAGTGGAGATCACGGCAACTGTCGTCTTGGAAGGACAACTTAAGGTATTGGG   719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1551 TGATGAATCTAGTGGAGATCACGGCAACTGTCGTCTTGGAAGGACAACTTAAGGTATTGGG   1610

Qy     720 AGATCGATGGGCAAAACTCTGTAGATGGACCAGGACCGCTGGGTCTTCTTTTACAAGACAT   779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1611 AGATCGATGGGCAAAACTCTGTAGATGGACCAGGACCGCTGGGTCTTCTTTTACAAGACAT   1670

Qy     780 CCTTCTCAAATGGCAAGGTCCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCATTCAGA   839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1671 CCTTCTCAAATGGCAAGGTCCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCATTCAGA   1730

Qy     840 AAAAGAAGATGCAGTGAACAAGATTCACCAACTGGCTTTAAAGATCAAAATGAAATGTT   899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1731 AAAAGAAGATGCAGTGAACAAGATTCACCAACTGGCTTTAAAGATCAAAATGAAATGTT   1790

Qy     900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGGATCTAGAAAGAAAAAGCAATCCAT   959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1791 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCCGGATCTAGAAAGAAAAAGCAATCCAT   1850

Qy     960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCACACTGAAGATAAGTCAGTGAC   1019
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1851 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCACACTGAAGATAAGTCAGTGAC   1910

Qy    1020 CCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTAGTCCAAAA   1079
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1911 CCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTAGTCCAAAA   1970

Qy    1080 ACTTGAAAAGAGTACAGACACAGACCCTTGAAAGA   1113
           ||||||||||||||||||||| |||||
Db    1971 ACTTGAAAAGAGTACAGCACAGGACGAAAACTGAAATA   2004
```

```
RESULT 3
AX114289
LOCUS          AX114289                5952 bp    DNA        linear     PAT 11-MAY-2001
DEFINITION     Sequence 1 from Patent WO0129943.
ACCESSION      AX114289
VERSION        AX114289.1  GI:14031259
KEYWORDS       .
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
  AUTHORS      Paul,X.L. and Xiao,X.
  TITLE        Method and vector for producing and transferring trans -spliced
               peptides
  JOURNAL      Patent: WO 0129243-A 1 26-APR-2001;
               DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
               SYSTEM OF HIGHER EDUCATION (US)
FEATURES             Location/Qualifiers
     source          1. .5952
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    2897. .2898
                     /note="S4 junction site"
     misc_feature    3198. .3199
                     /note="S2 junction site"
ORIGIN

Query Match          68.2%;     Score 1092.4;  DB 6;   Length 5952;
Best Local Similarity 99.4%;    Pred. No. 3e-257;
Matches 1107;  Conservative   0;  Mismatches   6;  Indels   6;  Gaps   1;

Qy       1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     891 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   950

Qy      61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTCATTGGCAGTCATTGATGGAGAG   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     891 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTCATTGATGGAGAG   1010

Qy     121 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1011 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC   1070

Qy     181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGGTGGTGAAAGA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1071 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGGTGGTGAAAGA   1130

Qy     241 CCAGTTTCATACTTCATGAGGGGTACAGCCCATCAGCGGCCGGGTTGG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1131 CCAGTTTCATACTTCATGAGGGGTACAGCCCATCAGCGGCCGGGTTGG   1190

Qy     301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1191 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA   1250
```

ORIGIN   /note="Synthetic"

Query Match 68.1%; Score 1091; DB 6; Length 8689;
Best Local Similarity 99.9%; Pred. No. 6.8e-257;
Matches 1102; Conservative 0; Mismatches 0; Indels 0; Gaps 1;

```
Qy    1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

Qy  121 TGAAGTAAACTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTTC 1278

Qy  181 TGCTGAGGACACATTGCAAGCACGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1279 TGCTGAGGACACATTGCAAGCACGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy  241 CCAGTTTCATACTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG   300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1339 CCAGTTTCATACTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy  301 TAATATTCTACAATTGGGAGTAAGCTGATTGGATTCAGAAGATGAAGA   360
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1399 TAATATTCTACAATTGGGAGTAAGCTGATTGGATTCAGAAGATGAAGA 1458

Qy  361 AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1459 AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy  421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATC-GAAACT   479
        |||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db 1519 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1578

Qy  480 GAAAGAGTTGAATGACTGGCTAACAGAGACAGAAGAGAACAAGGAAGA   539
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1579 GAAAGAGTTGAATGACTGGCTAACAGAGACAGAAGAGAACAAGGAAGA 1638

Qy  540 GCCTCTTGGACCTGATCTTGAAGACAACGCCAAGTACAACAACATAAGGTGCTTCA   599
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1639 GCCTCTTGGACCTGATCTTGAAGACAACGCCAAGTACAACAACATAAGGTGCTTCA 1698

Qy  600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACATGGTGGTGGTAGT   659
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACATGGTGGTGGTAGT 1758

Qy  660 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG   719
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1759 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy  720 AGATCGATGGGCAAACATCTGTAGATGGCAAACAGTGCTTTTAAAGATCAAAATGTT   779
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1819 AGATCGATGGGCAAACATCTGTAGATGGCAAACAGTGCTTTTAAAGATCAAAATGTT 1878

Qy  780 CCTTCCAAATGGCAACGTGGCCGTTTAAAAGCGGATCGAGAAAGAAAAGCAATCCAT   839
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1879 CCTTCCAAATGGCAACGTGGCCGTTTAAAAGCGGATCGAGAAAGAAAAGCAATCCAT 1938

Qy  840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT   899
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998

Qy  900 ATCAAGTCTTCAAAAACTGGCAACGTCGCCGTTTAAAAGCGGATCGATCTAGAAAAGCAATCCAT   959
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1999 ATCAAGTCTTCAAAAACTGGCAACGTCGCCGTTTAAAAGCGGATCGATCTAGAAAAGCAATCCAT 2058

Qy  960 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACCTGAAGAATAAGTCAGTGAC  1019
```

```
Qy  361 AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1251 AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1310

Qy  421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATC-GAAACT   479
        |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db 1311 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1370

Qy  480 GAAAGAGTTGAATGACTGGCTAACAGAGACAGAAGAGAACAAGGAAGA   539
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1371 GAAAGAGTTGAATGACTGGCTAACAGAGACAGAAGAGAACAAGGAAGA 1430

Qy  540 GCCTCTTGGACCTGATCTTGAAGACAACGCCAAGTACAACAACATAAGGTGCTTCA   599
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1431 GCCTCTTGGACCTGATCTTGAAGACAACGCCAAGTACAACAACATAAGGTGCTTCA 1490

Qy  600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACATGGTGGTGGTAGT   659
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1491 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACATGGTGGTGGTAGT 1550

Qy  660 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG   719
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1551 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1610

Qy  720 AGATCGATGGGCAAACATCTGTAGATGGCAAACAGTGCTTTTAAAGATCAAAATGTT   779
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1611 AGATCGATGGGCAAACATCTGTAGATGGCAAACAGTGCTTTTAAAGATCAAAATGTT 1670

Qy  780 CCTTCCAAATGGCAACGTGGCCGTTTAAAAGCGGATCGAGAAAGAAAAGCAATCCAT   839
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1671 CCTTCCAAATGGCAACGTGGCCGTTTAAAAGCGGATCGAGAAAGAAAAGCAATCCAT 1730

Qy  840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT   899
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1731 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1790

Qy  900 ATCAAGTCTTCAAAAACTGGCAACGTCGCCGTTTAAAAGCGGATCGATCTAGAAAAGCAATCCAT   959
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1791 ATCAAGTCTTCAAAAACTGGCAACGTCGCCGTTTAAAAGCGGATCGATCTAGAAAAGCAATCCAT 1850

Qy  960 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACCTGAAGAATAAGTCAGTGAC  1019
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1851 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACCTGAAGAATAAGTCAGTGAC 1910

Qy 1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGTGTTGGGATAATTTAGTCCAAA  1079
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1911 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCCGTGTTGGGATAATTTAGTCCAAAA 1970

Qy 1080 ACTTGAAAAGAGTACAGCACACAGACCCCTTGAAAGA 1113
        |||||||||||||||||||||||||||||||||||||
Db 1971 ACTTGAAAAGAGTACAGCACACAGACCCTGAAAATA 2004
```

```
RESULT 4
AX538622
LOCUS       AX538622      8689 bp   DNA    linear   PAT 23-NOV-2002
DEFINITION  Sequence 42 from Patent WO0229056.
ACCESSION   AX538622
VERSION     AX538622.1  GI:25271171
KEYWORDS    synthetic construct
SOURCE      synthetic construct
ORGANISM    artificial sequences
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 42 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES            Location/Qualifiers
     source         1..8689
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
```

```
Qy    540  GCCTCTTGGACCTGATCTTGAAGACCTAAAACCGCCAAGTACACAACAACATAAGGTGCTTCA  599
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639  GCCTCTTGGACCTGATCTTGAAGACCTAAAACCGCCAAGTACACAACAACATAAGGTGCTTCA  1698

Qy    600  AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCACTCACATCATGGTGGTGGTAGT  659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699  AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCCACTCACATCATGGTGGTGGTAGT  1758

Qy    660  TGATGAATCTAGTGAGATCACGGCAACTGCTGTGTTTGGAAGAACCAACTTAAGGTATTGGG  719
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759  TGATGAATCTAGTGAGATCACGGCAACTGCTGTGTTTGGAAGAACCAACTTAAGGTATTGGG  1818

Qy    720  AGATCGATGGGCAAACATCTGTAGTGGCAGAGAACGCGCTGGGTTCTTTTACAAGACAT  779
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819  AGATCGATGGGCAAACATCTGTAGTGGCAGAGAACGCGCTGGGTTCTTTTACAAGACAT  1878

Qy    780  CCTTCTCAAATGGCAACGTCTTACTGGAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879  CCTTCTCAAATGGCAACGTCTTACTGGAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  1938

Qy    840  AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAGATCAAAATGAAAATGTT  899
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939  AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAGATCAAAATGAAAATGTT  1998

Qy    900  ATCAAGTCTTCAAAAACTGCCGTTTAAAAGCCGATCTAGAAAAGAAAGCAATCCAT  959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999  ATCAAGTCTTCAAAAACTGCCGTTTAAAAGCCGATCTAGAAAAGAAAGCAATCCAT  2058

Qy    960  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  1019
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  2118

Qy   1020  CCAGAGACGCGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  1079
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119  CCAGAGACGCGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  2178

Qy   1080  ACTTGAAAAGAGTACAGCACAGA  1102
           |||||||||||||||||||||||
Db   2179  ACTTGAAAAGAGTACAGCACAGA  2201
```

RESULT 6
AX538627
LOCUS   AX538627   12057 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 47 from Patent WO0229056.
ACCESSION   AX538627
VERSION   AX538627.1   GI:25271181
KEYWORDS   .
SOURCE   synthetic construct
ORGANISM   synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS   Chamberlain,J.S. and Harper,S.Q.
TITLE   Mini-dystrophin nucleic acid and peptide sequences
JOURNAL   Patent: WO 0229056-A 47 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
source   1..12057
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="synthetic"
ORIGIN

Query Match   68.1%;   Score 1091;   DB 6;   Length 12057;
Best Local Similarity 99.9%;   Pred. No. 6.9e-257;
Matches 1102;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

```
Qy     1  GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099  GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  1158
```

```
Db   2059  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  2118

Qy   1020  CCAGAGACGCGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  1079
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119  CCAGAGACGCGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  2178

Qy   1080  ACTTGAAAAGAGTACAGCACAGA  1102
           |||||||||||||||||||||||
Db   2179  ACTTGAAAAGAGTACAGCACAGA  2201
```

RESULT 5
AX538624
LOCUS   AX538624   11443 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 44 from Patent WO0229056.
ACCESSION   AX538624
VERSION   AX538624.1   GI:25271175
KEYWORDS   .
SOURCE   synthetic construct
ORGANISM   synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS   Chamberlain,J.S. and Harper,S.Q.
TITLE   Mini-dystrophin nucleic acid and peptide sequences
JOURNAL   Patent: WO 0229056-A 44 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
source   1..11443
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Synthetic"
ORIGIN

Query Match   68.1%;   Score 1091;   DB 6;   Length 11443;
Best Local Similarity 99.9%;   Pred. No. 6.8e-257;
Matches 1102;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

```
Qy     1  GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099  GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  1158

Qy    61  TCCTTCACAGCATTGGAAGCTCCTGAAGCAAGTCATTGGCAGTTCATTGATGAGGAG  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159  TCCTTCACAGCATTGGAAGCTCCTGAAGCAAGTCATTGGCAGTTCATTGATGAGGAG  1218

Qy   121  TGAAGTAAACCTGGACCCGTTATCAAACAGCTTAGAAGAGTATTATCGTGGCTTCTTTC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219  TGAAGTAAACCTGGACCCGTTATCAAACAGCTTAGAAGAGTATTATCGTGGCTTCTTTC  1278

Qy   181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGGAAGTGGTGAAAGA  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGGAAGTGGTGAAAGA  1338

Qy   241  CCAGTTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGGCCGGGTTGG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339  CCAGTTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGGCCGGGTTGG  1398

Qy   301  TAATATTCTACAATTGGGAAGTAAGCTCAGACAGAAAATTATCAGAAGATGAAGA  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1399  TAATATTCTACAATTGGGAAGTAAGCTCAGACAGAAAATTATCAGAAGATGAAGA  1458

Qy   361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  1518

Qy   421  TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAAATC-GAAACT  479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  1519  TAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCAGAAACT  1578

Qy   480  GAAAAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGAAAATGGAGAAGA  539
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579  GAAAAGTTGAATGACTGGCTAACAAAACAGAGAAAGAACAAGAAAATGGAGAAGA  1638
```

RESULT 7
AX409637

LOCUS   AX409637   13957 bp   DNA   linear   PAT 14-JUN-2002
DEFINITION   Sequence 2284 from Patent WO0229103.
ACCESSION   AX409637
VERSION   AX409637.1 GI:21442342
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE   Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2284 11-APR-2002;
             GENE LOGIC INC (US)
FEATURES   Location/Qualifiers
     source   1..13957
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. M18533"
ORIGIN

Query Match   68.1%; Score 1091; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 6.9e-257;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy   1 GAGCTATGCTCACACCAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCTCACACCAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAGTTCATTGATGGGAGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAGTTCATTGATGGGAGA 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATCTCATGAGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTCATCTCATGAGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGAAGCTGATGGAACAGGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGAAGCTGATGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   421 TAGCATGGAAAACAAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   1519 TAGCATGGAAAACAAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAGAACAAGGAAAATGGAGGAGA 539
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAGAACAAGGAAAATGGAGGAGA 1638

Qy   540 GCCTCTTGGACCTGATCTTGAAGACTACAACAACATAAGGTGCTTCA 599
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACTACAACAACATAAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGAACAGTCAGGTCAGGTCAATTCTTCACTCACATGGTGGTAGT 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAGTCAGGTCAGGTCAATTCTTCACTCACATGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGGAGATCCACGCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG 719

```
Db   1759 TGATGAATCTAGTGGAGATCGAGAATCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGTGAAAGA 1338

Qy    720 AGATCGATGGGCAAACATCTGTAGATGACAGAGAACCGCTGGGTTCTTTTACAAGACAT 779
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 300

Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGACAT 1878
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy    780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 TAATATTCACAATTGGAAGTAAGCTGATTGGAACACGGAAAATTATCAGAGATGAAGA 360

Db   1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAGATGAAGA 1458

Qy    840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGATGGCCTCAGGGTAGC 420

Db   1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGATGGAATGCCTCAGGGTAGC 1518

Qy    900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAGAAAAGCAATCCAT 959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479

Db   1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAAAGAAAAGCAATCCAT 2058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy    960 GGGCAAACTGTATTCACTCAAACAAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGAGAAAATGAGGAGAA 539

Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC 2118
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGAGAAAATGAGGAGAA 1638

Qy   1020 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 1079
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    540 GCCTCTTGGACCTGATCTTGAAGACATAAACGCCAAGTACAACATAAGGTGCTTCA 599

Db   2119 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 2178
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACATAAACGCCAAGTACAACATAAGGTGCTTCA 1698

Qy   1080 ACTTGAAAAGAGTACAGCACAGA 1102
          |||||||||||||||||
Qy    600 AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 659

Db   2179 ACTTGAAAAGAGTACAGCACAGA 2201
          |||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1758

Qy    660 TGATGAATCTAGTGGAGATCGAGAATCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719

Db   1759 TGATGAATCTAGTGGAGATCGAGAATCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818
```

RESULT 8
X538581

```
LOCUS       AX538581               13957 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO0229056.
ACCESSION   AX538581
VERSION     AX538581.1  GI:25271086
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Chamberlain,J.S. and Harper,S.Q.
TITLE       Mini-dystrophin nucleic acid and peptide sequences
JOURNAL     Patent: WO 0229056-A 1 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

Query Match        68.1%;  Score 1091;  DB 6;  Length 13957;
Best Local Similarity  99.9%;  Pred. No. 6.9e-257;
Matches 1102;  Conservative 0;  Mismatches 0;  Indels 1;  Gaps 1;
```

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTATGCCACCACCTCTGACCCTACACGGAGCCCATT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTATGCCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGAACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGAACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACACAGCTTAGAAGACTTAGAAGAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACACAGCTTAGAAGACTTAGAAGAGTATTATCGTGGCTTCTTTC 1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGTGAAAGA 240
          |||||||||||||||||||||||||||||||
```

RESULT 9
HUMDYS

```
LOCUS       HUMDYS                 13957 bp    mRNA    linear   PRI 25-MAY-2000
DEFINITION  Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION   M18533 M17154 M18026 M20250
VERSION     M18533.1  GI:181856
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
             1 (bases 1 to 1699)
AUTHORS      Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and
             Kunkel,L.M.
TITLE        Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
             preliminary genomic organization of the DMD gene in normal and
             affected individuals
JOURNAL      Cell 50 (3), 509-517 (1987)
MEDLINE      87273512
PUBMED       3607877
REFERENCE    2 (bases 1678 to 3830)
AUTHORS      Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
TITLE        Conservation of the Duchenne muscular dystrophy gene in mice and
             humans
JOURNAL      Science 238 (4825), 347-350 (1987)
MEDLINE      88018015
PUBMED       3659917
REFERENCE    3 (bases 1 to 13957)
AUTHORS      Koenig,M., Monaco,A.P. and Kunkel,L.M.
TITLE        The complete sequence of dystrophin predicts a rod-shaped
             cytoskeletal protein
JOURNAL      Cell 53 (2), 219-226 (1988)
MEDLINE      88194521
PUBMED       3282674
COMMENT      On May 25, 2000 this sequence version replaced gi:340693.
             Draft entry and computer-readable sequence kindly provided by
             M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
             determined by the size of the deleted DNA segment. Deletions found
             in different patients were from positions 302-2200, 473-1168,
             1691-1810, and 1169-3011.
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xp21.3-p21.1"
                     /tissue_type="muscle"
                     /dev_stage="fetus"
                     /note="G00-119-850"
     gene            1..13957
                     /gene="DMD"
     CDS             209..11266
                     /gene="DMD"
                     /codon_start=1
                     /product="dystrophin"
                     /protein_id="AAA53189.1"
                     /db_xref="GI:181857"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSRTVHALNNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVKTKEEHFQLHHQMHYSQ
                     QITVSLAQGYERTSSPKPRFKSYAYTQAAVTTDSPATRTSPFFPSQHLEAPEDKSFGSSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMDLTAH
                     QGRVGNIQLGSKLIGTGKLSEDETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLKELNDWLTKTEERTRKMEEEPLGPDLEDLKRQVQQHKVLQEDLEQQVRVN
                     SLTHMVVVDESSGDHATALEEQLKVLGDRWANICRWTEDRWLLLQDILLKWQRLTE
                     EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLQKLAVLKADLEKKKQSMGKLYSLK
                     QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTTQPSLIQTTVM
                     ETVTVVTTREQILVKHAQEELPPPPQKKRQITVDSEIRKRLDVDITBLHSWITRSEA
                     VLQSPEFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSAQALVEQMVNEGVNA
                     DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIAFYNQLQLEQMTTAERNWLKIQ
                     PTTPSEPTAIKSQLKICKDEVNRLSGLQPQIERLKIQSALKEKGQGPMFLDADFVAF
                     TNHFKGVFSDVQAREKLQTIFDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
                     EIMEQRLGELQALQSSLQEQQSGLYYLSTTVKEMSKKAPSEISRKYQSEFEEIEGRWK
                     KLSSQLVEHCQKLEEQMNKLRKIQNHIQTIKKWWAEVDVFLKEEWPALGDSEILKKQL
                     KQCRLLVSDIQTIQPLNSVNEGGQKIKNEAEPFFASRIETELKELNQLDWRMCQQVY
                     ARKEALKGGLEKTVSLQKDLSEMHEWMTQAEEYLERDFEYKTPDELQKAVEEMKRAK
                     ESAQQEAKVKLLTESVNSVIAQAPPVAQBALKKELEITLTTNYQWLCTRLNGKCKTLE
                     EVWACWHELLSYLEKANKWLNEVEFKLKTENIPGGAEISEVLDSLENLMRHSEDNP
                     NQIRILAQTLTDGGVMDELINEELETFNSRWRELHEEAVRRQKLLEQSIQSAQETEKS
                     LHLIQESLTFIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLSEEMKKHQGKEA
```

```
AQRVLSQIDVAQKKLQDVSMKFRLFQKPANFELRLQESKMILDEVKMHLPALETKSVE
QEVVQSQLNHCVNLYLKSLSEVKSEVEMVIKTGRQIVQKKQTENPKELEDRVTALKLHY
NELGAKVTERKQOLEKCLKLSRKMRKEMNVLTEWLAATCMELTKRSAVEGMPSNLDSE
VAWGKATQKEIEKQVHLKSITEVGEALKTVLGKKETLVEDKLSLLNSNWIAVTSRAE
EWLNLLLEYQKHMETFDQNVDHITKWIIQADTLDSEKKKPQQEDKVLRLKALEND
IRPKVDSTRDQAANLMANRGDHCRKLVEPQISELNHRFAAISHRIKTGKASIPLKELE
QFNSDIQKLLEPLEABIQQVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQQRITDER
KREEIKIKQQLLQTKHNALKDLRSQRRKALEISHQWYQKQADDLLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKKBELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWRE
IESKPAQFDRLNFAQIHTVREETMVVMTEDMPLEISYVPSTYLTEITHVSQALLVEKL
LLNAPLCAKDFEDLFKQEESLKNIKDSLQQSGRIDIIHSKKTAALQSATPERVKL
QEALSQLDFQWEKVNKMYKDRQGRFDRSVEKWRRFHYDIKIFNQWLTEABQFLRKTQI
PENWEHAKYKWYLKELQDGIGQRQTVVRTLNATGEEIIQQSSKTDASILQEKLGSLNL
RWQEVCKQLSDRKKRLEEQKNILSEFQRDLNEFVIWLBEADNIASIPLEPGKEQQLKE
KLEQVKLLVBELPLRQGIIKQLNETGGPVLVSAPISPEEQDKLENKLKQTNLQWIKVS
RALPEKQGEIEAQIKDLGQLEKKLEDLEEQLNHLLLWLSPIRNQLEIYNQPNOEGPFD
VQETEIAVQAKQPDVEBILSKGQHLYKEKPATQPVKRKLEDLSSEWKAVNRLLQELRA
KQPDLAPGLTTIGASPTQYVTLVTQPVVTKETAISKLEMPSSLMLEVPALADPNRAWT
ELTDWLSLLDQVIKSQRVMVGDLEDINEMIIKQRATMODLEQRRPQLELELITAAQNLK
NKTSNQEARTIIIDRIERIQNQWDEVQELHLQNRRQQLNEMLKDSTOWLEAKEEAEQVL
GQARAKLESWKEGPYTVDAIQKKITETKQLAKDLRQWCTNVDVANDLAIKLLRDYSAD
DTRKVHMITENINASWRSIHKRVSREAALEETHIRLLQQFPLDIEKFLAWLTEATTA
NVLQDATRKERLLEDSKGVKELMKQWQDLQGEIEAHTDVYHNLDENSOKILRSLEGSD
DAVLIQRRLRDMNNFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELLVWLQLKDDEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKTKEPYVMSTLBTVRIFLTEQPLEGLEKLYQ
EPRELPPEERAQNVTRLLRKQAEBVNTEWKLNLHSADWQRKIDETLERLQELQEATD
ELDLKLRQAEVIKSWQPVGDLLIDLSLQDHLEKVKALRGEIAPLKENVSHVNDLARQL
TTLGIQLSPYNLSTLEDLNTRWKLLQVAVEHYGRIBEERGBLERIIADLEBENRNLQAEY
WERAISPNKVPYYINHETQTTCWDHPKWTELYQSLADLNVRFSAYRTAMKLRRLQKA
LCLDLLSLSAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHNNLVNVDTGT
MCLNWLNVDTGRTGRIRVLSFKFTGIIISLCKAHLEDKYRYLFKQVASSTGFCDQRRL
GLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIEAALFLDWMRLEPQSM
VNLPVLHRVAAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDICQSCFFSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVCTVLEGDNMET
PVTLINFWPVDSAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSYLNDSISPNESID
DEHLLIHYCQSLNQDSPLSQPRSPAQIILSLESEERGBLERIIADLEBENRNLQAEY
DRLKQQHEHKGLSPLPSPPEMPTSPQSPRDAELIABEAKLLROHKGRLEARMOILEDH
NKQLESQLHRLRQLLEQPQABAKVNGTTVSSPSTSLQRSDSSQPMLLRVVGSQTSDSM
GEEDLLSPPQDTSTGLEEVMEQLNNSFPPSRGRNTPGKPMREDTM"
```

```
ORIGIN
          Query Match              68.1%;    Score 1091;  DB 9;   Length 13957;
          Best Local Similarity    99.9%;    Pred. No. 6.9e-257;
          Matches  1102;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  1;

Qy     1  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  1158

Qy    61  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGATGGAGAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGATGGAGAG  1218

Qy   121  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAGTATTATCGTGGCTTCTTTC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAGTATTATCGTGGCTTCTTTC  1278

Qy   181  TGCTGAGGACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  1338

Qy   241  CCAGTTCATACTCATGAGGGTACATGAGATTGACAGCCCATCAGGGCCGGGTTGG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339  CCAGTTTCATACTCATGAGGGTACATGAGATTGACAGCCCATCAGGGCCGGGTTGG  1398

Qy   301  TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1399  TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA  1458

Qy   361  AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGAATGCTCAGGGTAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCTCAGGGTAGC  1518
```

```
Qy      1 GAGCTATGCCTACACACAGCCTGCTTATGTCACCACCTTCTGACCCTACACGCGAGCCCATT    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGCCTGCTTATGTCACCACCTTCTGACCCTACACGCGAGCCCATT  1158

Qy     61 TCCTTCACAGCATTGGAGCTCCTGAAGACAAGTCATTGGAGTTCATTGATGGGAGAG   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAGCTCCTGAAGACAAGTCAGTCATTGGCAGTTCATTGATGATGGGAGAG  1218

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTGCTTCTTTC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTGCTTCTTTC  1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGATTCTAATGATGTGGAAGTGGTGAAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  1338

Qy    241 CCAGTTTCATCATGAGGGGTACATCAGCCCATCAGCGCCGGTTGG   300
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATCATGAGGGGTACATCAGCCCATCAGCGCCGGTTGG  1398

Qy    301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA  1458

Qy    361 AACTGAAGTACAAGACGATGATCTCCTAAATTCAAGATGGAATGGCCTCAGGGTAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGACGATGATCTCCTAAATTCAAGATGGAATGGCCTCAGGGTAGC  1518

Qy    421 TAGCATGGAAAACAAAGCAATTTCACATAGAGTTTTAATGGAATCTCCAGAATC-GAAACT   479
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||| ||||||
Db   1519 TAGCATGGAAAACAAAGCAATTTCACATAGAGTTTTAATGGAATCTCCAGAATCAGAAACT  1578

Qy    480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAAGAACAAGGAAAATGAGAGGAAGA   539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAAGAACAAGGAAAATGAGAGGAAGA  1638

Qy    540 GCCTCTTGACCTGATCTTGAAGACCAAGCAAGTCAATTCTCTCACTCACATAAGGTGCTTCA   599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGACCTGATCTTGAAGACCAAGCAAGTCAATTCTCTCACTCACATAAGGTGCTTCA  1698

Qy    600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATCTGCTTGGAAGAACAACTCAAGTATTGGG   659
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATCTGCTTGGAAGAACAACTCAAGTATTGGG  1758

Qy    660 TGATGAATCTAGTGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAAGACAT   719
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGCAAACATCTGTAGATGACAGAAGACCGCTGGGTTCTTTTACAAGACAT  1818

Qy    720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGGACCAGAGGACCTGGGTTCTTTTAGTGCATGGCTTTCAGA   779
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAGGACCAGAGGACCTGGGTTCTTTTAGTGCATGGCTTTCAGA  1878

Qy    780 CCTTCTCAAATGGCAACGTGTTACTGGACTTGAAGAACAACAGTGCCTTTTTAGTGCATGGCTTTCAGA   839
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTGTTACTGGACTTGAAGAACAACAGTGCCTTTTTAGTGCATGGCTTTCAGA  1938

Qy    840 AAAGAAGATGCAGTGAACAGATTCACACAACTGGCTTTAAAGATCAAAATGAAAATGTT   899
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAGAAGATGCAGTGAACAGATTCACACAACTGGCTTTAAAGATCAAAATGAAAATGTT  1998

Qy    900 ATCAAGTCTTCAAAAACTGCCGTTTAAAGGCGGATCTAGAAAGACGAAAAGCAATCCAT   959
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGCCGTTTAAAGGCGGATCTAGAAAGACGAAAAGCAATCCAT  2058

Qy    960 GGGCAACTGTATTCTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  1019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAACTGTATTCTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC  2118

Qy   1020 CCAGAAGACCGAAGCAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  1079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAAGACCGAAGCAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA  2178

Qy   1080 ACTTGAAAAGAGTACAGCACAGA  1102
          |||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGA  2201
```

ESULT 10
R220819
OCUS   AR220819   13977 bp   DNA   linear   PAT 26-SEP-2002
EFINITION   Sequence 60 from patent US 6426186.
CCESSION   AR220819
ERSION   AR220819.1   GI:23327696
EYWORDS
OURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
EFERENCE   1 (bases 1 to 13977)
  AUTHORS   Jones,K.A., Volkmuth,W. and Walker,M.G.
  TITLE   Bone remodeling genes
  JOURNAL   Patent: US 6426186-A 60 30-JUL-2002;
EATURES   Location/Qualifiers
  source   1..13977
                     /organism="unknown"
                     /mol_type="genomic DNA"
RIGIN

Query Match   68.1%; Score 1091; DB 6; Length 13977;
Best Local Similarity   99.9%; Pred. No. 6.9e-257;
Matches 1102; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

```
Db  2179 ACTTGAAAAGAGTACAGCACAGA 2201
         |||||||||||||||||||||||
```

```
RESULT 11
LOCUS       HSDMDR
DEFINITION  Human mRNA for dystrophin.
                        12446 bp    mRNA    linear   PRI 12-SEP-1993
ACCESSION   X14298
VERSION     X14298.1  GI:30845
KEYWORDS    Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 12446)
  AUTHORS   Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
            Davies,K.E.
  TITLE     Two human cDNA molecules coding for the Duchenne muscular dystrophy
            (DMD) locus are highly homologous
  JOURNAL   Nucleic Acids Res. 17 (13), 5391 (1989)
  MEDLINE   89345106
   PUBMED   2668885
REFERENCE   2  (bases 1 to 12446)
  AUTHORS   Rosenthal,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
            der DDr, Zentralinstitut fuer Molekularbiologie, Robert-Roessle
            Str.10, 1115 Berlin Buch, DDR
COMMENT     see also M18533 and M20250 for Dmd segs.; discrepancies compared to
            M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449,
            3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES             Location/Qualifiers
     source          1..12446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X chromosomal, Xp21."
                     /tissue_type="muscle"
                     /dev_stage="fetal and adult."
     CDS             99..11156
                     /note="unnamed protein product; dystrophin (AA 1 - 3685)"
                     /codon_start=1
                     /protein_id="CAA32479.1"
                     /db_xref="GI:30846"
                     /db_xref="GOA:P11532"
                     /db_xref="SWISS-PROT:P11532"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNNVNKALRVLQNNVDLVNIGSTDIVDG
                     NHKLTIGLIWNIILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVIKEEHFQLHHQMHYSQ
                     QITVSLAQGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFPSQHLEAPEDKSFGSSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMDLTAH
                     QGRVGNILQLGSKLIGTGKLSEDEETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLKELNDWLTKTEERTRKMEEPLGPDLEDLKRWANICRWTEDRWVLLQDILLKWQRLTE
                     EQCLFSAWLSKEDAVNKIHTTGFKDQNEMLSSLQKLAVLKADLEKKKQSMGKLYSIK
                     QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTTQPSLTQTTVM
                     ETVTTVTREQLVKHAQEELPPPPQKKRQITVDSEIRKRLDVDITELHSWITRSEA
                     VLQSPEFAIFRKEGNFSDLKEKVNAIEREKAEKPRKLQDASRSGQALVEQMVNEGVNA
                     DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLEQMTTAENWLKIQ
                     PTTPSEPTAIKSQLKICKDEVNRLSGLQPQIERLKIQSIALKEKGQGPMFLDADFVAF
                     TNHFKQVFSDVQAREKSLQTIFDTLPPMRYQETMSAIRTWVQQSETKLSSIPQLSVTDY
                     EIMEQRLGELQALQSSLQEQSGLYLYLSTVKEMSKKAPSEISRKYQSEFEEIEGRWK
                     KLSSQLVEHCQKLEEQMNKLRKIQNHIQTLKKWMAEVDVFLKEEWPALGDSEILKKQL
                     KQCRLLVSDIQTIQPSLNSVNEGGQKIKNEAEPEFASRLETELKELNTQWDHMCQQVY
                     ARKEALKGGLEKTVSLQKDLSEMHEWMTQAEEYLERDFEYKTPDELQKAFEMKRAK
                     EBAQQKEAKVKLLTESVNSVAQPVAQPVAQALKKELETLITTNYQWLCTRLNGKCKTLE
                     EVWACWHELLSYLEKANKWLNEVEFKLKTTENIPGGAEEISEVLDSLENLMRHSEDNP
                     NQIRLAQTLTDGVWDELINEBLETFNSRWRELHEEAVRRQLKLEQSIQSAQETENS
                     LHLIQESLTFIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLEMKKHNQGKEA
                     AQRVLSQIDVAQKLQDVSMKRLFQKPANFEQRIQESKMILDEVKMHLPALETKSVE
```

```
QEVVQSQLNHCVNLYKSLSEVKSEVEMVIKTGRQIVQKKQTENPKELDERVTALKLHY
NELGAKVTERKQLCLKLSRMRKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
VAWGKATQKEIBKQKVHLKSITEVGEALKTVLGKKETLVEDKLSLLNSNWIAVTSRAE
EWLNLLLEYQKHMETFDQNVDHITKWIIQADTLLDDSEKKKPQQKEDVLKRLKAELND
IRPKVDSTRDQAANLMANHGDHCRKLVEPQISELNHRFAAISHRIKTQGALSIPLKELE
QPNSDIQKLLEPLEAEIQQGVNLKEEDNKDMNEDNEGTVKELLQRGDNLQQRITDER
KSEEIKIKQLLQTYKNALKDLRSQRRKKALEISHQWYQYKRQADDLLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKKELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWRE
IESKFAQFRRLNFAQIHTVREETMMVMTEDMPLEISVVPSTYLTEITHVSQALLEVEQ
LLNAPDLCAKDFEDLFKQESLKNIKDSLQQSSGRIDIIHSKKTAALQSATPVERVKL
QEALSQLDFQWEKVNKMYKDRQGRFDRSVEKWRRFHYDIKIFNQWLTEAEQFLRKTQI
PENWEHAKYKWYLKELQDGIGQRQTVRTLNATGEBIIQQSSKTDASILQEKLGSLNL
RWQEVCKQLSDRKKRLEEQKNISEFQRDLNEFVLWLEEADNIASIPLEPGKEQQLKE
KLEQVLLVEBLPLRQGILKQLNETGGPYLVSAPISPEQDKLENKLKQTNLQWIKVS
RALPEKQGEIEAQIKDLGQLEKKLEDLEEQLNHLLLWLSPIRNQLEIYNQPNQEGPFD
VKETEIAVQAKQPDVEEILSKGQHLYEKPATQPVKRKLEDLSSEWKAVNRLLQELRA
KQPDLAPGLJTTIGASPTQVTLVTQPVVTKETAISKLKEMPSSLMEVPALADFNRAWT
ELTDWLSLLDQVIKSQRVMVGDLEDINEMIIKQKATMQDLEQRPQLEBLITAAQNLK
NKTSNQEARIIITDRIERIQNQWDEVQEHLQNREQLNEMLKDSTQWLEAKEBAEQVL
GQARAKLESWKEGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSAD
DTRKVHMITENINASWRSIHKRVSEREAALEETHRLLQQFPLDLEKFLAWLTEAETTA
NVLQDATRKERLLEDSKGVKELMKQWQDLQGEIEAHTDVYHNLDENSQKIRLSEGSD
DAVLJQRRLDMNNFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQEBLLVWLQKDDEL
SRQAPIGGDFPPAVCQKQNDVHRAFKRELKTKEPVIMSTLETVRIFLITEQPLEGLEKLYQ
EPRELPPEBRAQNVTRLLRKQAEEVNTEWEKLNLHSADWQRKIDETLETRLQELQEATD
ELDLKLRQAEVIKSWQPVGDLLLDSLQDHLEKVKALRGEIAPLKENVSHVNDLARQL
TTLGIQLSPYNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVQGP
WERAISPNKVPYYINHETQTCWDHPKMTELYQGLADLNNVRTGLKKLQEQ
LCLDLLSLSAACDALDQHNLKQNDQPMDILQINCLTTIYDRLEQEHNNLNVPLCVD
MCLNWLLNVDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRL
GLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIEAALFLDWMRLBPQSM
VNLPVLHRVAAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDICQSCFFSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLEGNMET
PVTLINFWPVDSAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSYLNDSISPNESID
DEHLLIQHYCQSLNQDSPJ.SQPRSPAQILISLSEERGBLERIIADLEENRNLQAEY
DRLKQQHEHKGLSPLPSPPEMMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDH
NKQLBSQLHRLRQLLEQPQAEAKVNGTTVSSPSTLSQRSDSSQPMLRVVGSQTSDSM
GEEDLLSPPQDTSTGLEEVMEQLNNSPFSSRGRNTPGKPMREDTM"
```

```
ORIGIN

Query Match     67.9%;   Score 1087.8;   DB 9;   Length 12446;
Best Local Similarity   99.7%;   Pred. No. 4.2e-256;
Matches 1100;   Conservative   0;   Mismatches   2;   Indels   2;   Gaps   1;

Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGCACCGAGCCCATT    60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   989 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGCACCGAGCCCATT  1048

Qy    61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCTTAGAAGAAGTATATCGTCATTGGTGGCTCATTGATGAGAG   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1049 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCTTAGAAGAAGTATATCGTCATTGGTGGCTCATTGATGAGAG  1108

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATATCGTGGCTCTTTC   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1109 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATATCGTGGCTCTTTC  1168

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGAGTGTGAAAGA   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1169 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGAGTGTGAAAGA  1228

Qy   241 CCAGTTCATACTCATCAGAGGGTACAGCAGCCCATCGACAGCCCATTGACAGCGCCGGTTGG   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1229 CCAGTTCATACTCATCAGAGGGTACAGCAGCCCATCGACAGCCCATTGACAGCGCCGGTTGG  1288

Qy   301 TAATATTCTACAATTGGGACGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1289 TAATATTCTACAATTGGGACGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  1348

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1349 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC  1408

Qy   421 TAGGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATGGATCTCCAGATC-GAAACT  479
```

```
FEATURES             Location/Qualifiers
     source          1..4402
                     /organism='Unidentified'.
FT   source          1..4402
FT                   /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

ORIGIN

Query Match    58.2%;  Score 932;  DB 6;  Length 4402;
Best Local Similarity  79.7%;  Pred. No. 7.4e-218;
Matches 1277;  Conservative  0;  Mismatches  0;  Indels  325;  Gaps  2;

Qy     1 GAGCTATGCCTACACACAGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCGCCCATT   60
           |||||||||||||||||||||||||||||||||||||||||||||||  ||  |||||||
Db  1095 GGCTATGCCTACACACAGCTGCTATGTCACCACCTCTGACCCTACACGCGAGCCGAGCCCATT 1154

Qy    61 TCCTTCACAGCATTGGAGCTCCTGAAGACAAGCAGTCATTGGCAGTTCATTGATGGAGAG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1155 TCCTTCACAGCATTGGAGCTCCTGAAGACAAGCAGTCATTGGCAGTTCATTGATGGAGAG 1214

Qy   121 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAAGAAGCAGTTATTATCGTGGCTTCTTTC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1215 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAAGAAGCAGTTATTATCGTGGCTTCTTTC 1274

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCAATGATGTGGAAGTGGTGAAAGA  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1275 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCAATGATGTGGAAGTGGTGAAAGA 1334

Qy   241 CCAGTTCATACTCATGAGGGGTACATGATTGACAGCCATCAGGGCCGGGTGG  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1335 CCAGTTCATACTCATGAGGGGTACATGATTGACAGCCATCAGGGCCGGGTGG 1394

Qy   301 TAATATTCTACAATTGGGAGTAAGTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1395 TAATATTCTACAATTGGGAGTAAGTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1454

Qy   361 AACTGAAGTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGGGAATGGGAATGGAGTAGC  420
Db  1455 AACTGAAGTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGGGAATGGGAATGGAGTAGC 1514

Qy   421 TAGCATGGAAAAACAAGCAATTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT  479
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  1515 TAGCATGGAAAAACAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1574

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAAGAACAAGGAAAATGAGGAGAA  539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1575 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAAGAACAAGGAAAATGAGGAGAA 1634

Qy   540 GCCTCTTGGACCTGATCTTGAAGACTCAAAACGCCAAGTACAACATAAGGTGCTTCA  599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1635 GCCTCTTGGACCTGATCTTGAAGACTCAAAACGCCAAGTACAACATAAGGTGCTTCA 1694

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTGGTAGT  659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1695 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTGGTAGT 1754

Qy   660 TGATGAATCTAGTGAATGACAGTCACGCGGATGGAGAACAACTTAAGGTATTGGG  719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1755 TGATGAATCTAGTGAATGACAGTCACGCGGATGGAGAACAACTTAAGGTATTGGG 1814

Qy   720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTGGTTCTTTTACAAGACAT  779
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1815 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTGGTTCTTTTACAAGACAT 1874

Qy   780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1875 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1934

Qy   840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT  899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1935 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1994
```

b 1409 TAGCATGGAAAAACAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1468
y   480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAAGAACAAGGAAAATGAGGAGAA 539
b 1469 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAAGAACAAGGAAAATGAGGAGAA 1528
y   540 GCCTCTTGGACCTGATCTTGAAGACTCAAAACGCCAAGTACAACATAAGGTGCTTCA 599
b 1529 GCCTCTTGGACCTGATCTTGAAGACTCAAAACGCCAAGTACAACATAAGGTGCTTCA 1588
y   600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTGGTAGT 659
b 1589 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTGGTAGT 1648
y   660 TGATGAATCTAGTGAATGACAGTCACGCGGATGGAGAACAACTTAAGGTATTGGG 719
b 1649 TGATGAATCTAGTGAATGACAGTCACGCGGATGGAGAACAACTTAAGGTATTGGG 1708
y   720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTGGTTCTTTTACAAGACAT 779
b 1709 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGTGGTTCTTTTACAAGACAT 1768
y   780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
b 1769 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1828
y   840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 899
b 1829 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1888
y   900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCCGATCAGAAAAGAAAAAGCAATCCAT 959
b 1889 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCCGATCAGAAAAGAAAAAGCAATCCAT 1948
y   960 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTCGAAGAATAAGTCAGTGAC 1019
b 1949 GGGCAAACTGTATTCACTCAAACAGATCTTCTTCAACACTCGAAGAATAAGTCAGTGAC 2008
y 1020 CCAGAGACCGAAGCATGGCGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
b 2009 CCAGAGACCGAAGCATGGCGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2068
y 1080 ACTTGAAAAGAGTACAGCACAGA 1102
b 2069 ACTTGAAAAGAGTACAGCACAGA 2091

```
Qy   900  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT  959
          |||||||
Db  1995  ATCAAGT----------------------------------------------------  2001

Qy   960  GGGCAAACTGTATTCACTCAACAAGATCTCTTTCACACTGAGAGAATAAGTCAGTGAC  1019
Db  2002  ----------------------------------------------------------  2002

Qy  1020  CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA  1079
Db  2002  ----------------------------------------------------------  2002

Qy  1080  ACTTGAAAAGAGAGTACAGCACACAGACCCCTGAAAGACTTCAAGAGGCCACGGA  1139
Db  2002  ----------------------------------------------------------  2002

Qy  1140  TGAGCTGGACCTCAAGCTCGCGCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG  1199
Db  2002  ----------------------------------------------------------  2002

Qy  1200  CGATCTCCTCATGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGGCACTTCGAGGAGA  1259
          ||||||||||||||||||||||||||||
Db  2002  ----------------------------CTCGAAAGTCAAGGCACTTCGAGGAGA  2030

Qy  1260  AATTGCGCCTCTCGAAAGAGAGAAGAACGTGAGGCCACGTCAATGACCTGCTCCGCCAGCTTACCAC  1319
          ||||||||||||||||||||||||||||||||||||||
Db  2031  AATTGCGCCTCTCGAAAGAGAGAAGAACGTGAGCCACGTCAATGACCTGCTCCGCCAGCTTACCAC  2090

Qy  1320  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTCGAAGACCTGAACACCAGATG  1379
          |||||||||||||||||||||||||||||||||||||
Db  2091  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTCGGAGACCTGAACACCAGATG  2150

Qy  1380  GAAGCTTCTGCAGGTGGCCGTCGAGAACTGGACCGAGTCAGGCAGCTGCATGAAGGCCCACGGGA  1439
          ||||||||||||||||||||||||||||||||||||||
Db  2151  GAAGCTTCTGCAGGTGGCCGTCGAGAACTGGACCGAGTCAGGCAGCTGCATGAAGGCCCACGGGA  2210

Qy  1440  CTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCGTCTGTCCAGGGTCCCTCGGGAGAGAGC  1499
          ||||||||||||||||||||||||||||||||||||||
Db  2211  CTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCGTCTGTCCAGGGTCCCTCGGGAGAGAGC  2270

Qy  1500  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACCACGAGACTCAAACAACTTGCTGGGA  1559
          ||||||||||||||||||||||||||||||||||||||
Db  2271  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACCACGAGACTCAAACAACTTGCTGGGA  2330

Qy  1560  CCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCT  1601
          |||||||||||||||||||||||||||
Db  2331  CCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCT  2372
```

FEATURES   Location/Qualifiers
     source   1..13887
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /strain="breed Golden retriever"
                     /db_xref="taxon:9615"
     5'UTR   <1..289
     CDS   290..11332
                     /codon_start=1
                     /product="dystrophin"
                     /protein_id="AAC83646.1"
                     /db_xref="GI:3982751"
                     /translation="MLWEVEDCYEREDVQKKTFTKVVNAQFSKFGKGHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPEKGSTRVHALNVNKALRVLQONNVDLVNIGSTDIVDG
                     NHKLTIGLIWNIIIHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCCQSATORLEHAFNIAKYQLGIEKLLDPEDV
                     ATTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPSQVTREEHFQIHHQMHYSQ
                     QITVSLAQGYERAPSFPKPRFKSYAYTQAAYVTTSDPTRSPLPSQHLETPEDKSFGRS
                     LTETEANLDSYQTALEEVLSWLLSAEDALQAAQGEISNDVEEVKEQFHTHEGYMMDLTS
                     HQGRVGNVLQLGSQLIGTGKLSEDEETEVQEOMNLLNSRWECLRVASMEKQSNLHKVL
                     MDLQNQQLKELNDWLTKTERTRMEKEPLGPDIEDLKRQVQQHKVLQEDLEQEQVRV
                     NSLTHKVVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRRVLLQDILLKWQRFT
                     EEQCLFSAWLSEKEDAVNKIHTTGFKDQSEVLSNLQKLAVLKTDLEKKQTMDKLCSL
                     NQDLLSALKNTVVAHKMEAWLDNFAQRWDNLVQKLEKSSAQISOAVTTTQPSLIQTTV
                     METVMVTTREHILVKHAQELPPPPQQKKRQIIVDSEIRKRLLDVDITELHSWITRSE
                     AVLQSPEFAIYRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSAQALVEQMVNEGVN
                     ADSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIITFNQLQLEQMTTAENWLKT
                     QPTTTSEPTAIKSQLKICKDEINRLSALQPIERLKIQSIALKEKGQGPMFLDADFVA
                     FTNHFNQVFADVQAREKELQTIFDSLPPMRYQETMSTILTWIQSETKLSIPQVTVTE
                     YDIMEQRLGELQALQSSLQEQQNGLNYLSTTVKEMSKKAPLSDISRKYQSEFEIEGR
                     WKKLSSQLVEBHCQKLEEOMAKLRKIQNHIKTLKKWITEVDVPLKEEWPALGDSEILKR
                     QLKQCRLLVNDIQTIQPSLNSVNSGAQKMKNEAEPEFAGRLETELRELNTQWDYMCRQ
                     VYARKEALKGGLDKTVSLQKDLSEMHEWMTQAEEYLERDFEYKTPDELQTAVEEMKR
                     AKEBEAQQKERAKVKLLTESVNSVIAQAPPAAQEALKKELDTLTNYQWLCTRLNGKCKT
                     LEEWVACWHELLSVIEKANKWLSEVEVKLKTTENISGGAEBIAEVLDSLENLMQHSED
                     NPNQIRLAQTLTDGGVMDELINELETNSRWRELHEBAVRKVLLEQSIQSAQEIE
                     KSLHLIQESLSSIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLEEMKKHNQGK
                     ETAQRVLSQIDVAQKKLQDVSMKFRLFQKPANFFQRLQESRMLQDERVTALKL
                     VEQEVVQSQLAHCVNLYKSLSEVKSEVEMVIKTGEQRIVQKKQTENPKELDERVTALKL
                     HYNELGAKVTERKQQLEKCLKLSRKMRKEMNALTEWLAATDMELTKRSAVEGMPSNLD
                     SEVAWGKATQKEIEKQKVHLKSVTEVGEALKTVLGKKEMLVEDILKRLKAEM
                     ABEWLNLLLEYQKHMETFDQNVDYITNWIIQADALLDEBEKKKPQQKEDILKRLKAEM
                     NDIRPKVDSTRDQAANLMANRGDHCRKVVBPKISELNHRFAAISHRIKTGKASIPLKE
                     LEQFNSDIQKLLEPLEAEIQGVNLKEEDFNKDMSEDNEGTVKELLQRGDNLQQRITD
                     ERKREIKIKQQLLQTKHNALKDLRSQRRKKALEISHQWYQYKRQADDLLKCLDDIEK
                     KLASLPEPRDERKIKEIDRELQKKEEBLNAVRRQAEGLSEDGAAMAVEPTQIQLSKRW
                     REIESKFAQFRRLNFAQIHTVHEBSVVAWTEDMPLEISVPPSTYLTEITHVSQALSEV
                     EELLNAPDLCAQDFEDLFKQEESLKNIKDSLQQISGRIDIIHNKKTAALHSATPABRA
                     KLQEALSRLDFQWERVNNMYKDRQQRFDRSVEKWRRFHYDMKIINQWLTEAEQFLKKT
                     QIPENWEHAKYKWYLKELQDGIGQRQSVVNLNATGEBIIQQSKTDASILQEKLGSL
                     NLRWQEVCKQLAERRKKRLEEQKNILSEFQRDVNEFVLWLEEADNVANIPLEPGNEQQL
                     KEKELEQVKLLAEBLPLRQGILKQLNETGGTVLVSAPLSPEEQDKLENKLKQTNLQWIK
                     VSRNLPEKQSEIEAVHVLDLGQLEEQLNHILLWLSPIRNQLEIYNQPNQTGPFDKEIE
                     VAVQAKQPDVEGILSKGQHLYKEKPATOPAKRKLEDLSSDWKVVTQLLQELRAKQPGP
                     APGLITVRAPPSQTVTLVTQPAVTKETAISKLEMPSSLLLEVPALADFNRAWTELTDW
                     LSLLDRVIKSQRVMVGDLEDINEMIKQKATLQDLEQRPQLELIITAAQNLKNKTSN
                     QEARTIITDRIERIQSOWDEVQEHLQNRLQLTEMLKDSTOWLEAKERABQVLGOARA
                     KLESWKEAPYTVDAIQKKTETKQLAKDLRQWINVDVANDLALKLLRDYSADDTRKV
                     HMITENNASWASIHKRLSEREALEETHRLLQQFPLDLEKFLAWLTEAETTANVLQD
                     ATHKERLLEDSKGVRELMKQWQDLQGEIEAHTDIYHNLDENGQKVLRSLEGSDDAALL
                     QRRLDNMNFKWSELRRKSLNIRSHLEASSDOWKRLHLSLQELLVWLQLKDDELSRQAP
                     IGGDFPAVQKQNDVHRAFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLYQEPREL
                     PEEERAQNVTRLLRKQAEBVNTQWEKLNVHSADWQRKIDEALERLQELQEATDELDLK
                     LRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEITPLKENVSYVNDLARQJTTLGI
                     QLSPYNINTLBDLNTRWKLLQVAIEDRIRQLHEAHRDFGPASQHFLSTSVQGPWERAI
                     SPNKVPYINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKALCLDL
                     LSLSAACDALDQHNLKQNDQPMDILQVINCLTTIYDRLEQEHNNLVNVPLCVDMCLNW
                     LLNVDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRLGLLLH
                     DSIQIPRQLGEVASFGGSNIBPSVRSCFQFANNKPEIEAALFLDWMRLEPQSMWLPV
                     LHRVAAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDICQSCFSGRVAKGHKMHYP
                     MVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLEGDNMETPVTLI
                     IQHYWRSLNQBSPLSQPRSPAIQILISLESEBERGLERIIADLEGRNRNIQAEYDRLKQ
                     QHEHKGLSPLPSPPEMMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDHNKQLE
                     SQLHRLRQLLEQPQABAKVNGTTVSSPSTSLQRSDSSQPMLLRVVGSQTSESMGEEDL
                     LSPPQDTSTGLEEVMEQLNHSFPSSRGRNTPGKPMREDTM"
     3'UTR   11333..13887
     repeat_region   11789..11828
                     /rpt_type=tandem
                     /rpt_unit="ca"
ORIGIN

Query Match   57.3%;   Score 916.6; DB 4; Length 13887;
Best Local Similarity 90.0%;   Pred: No. 4.8e-214;
Matches 993;   Conservative   0;   Mismatches 109;   Indels   1;   Gaps   1;

```
y      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1183 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCACACGGAGCCCACT 1242

y     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAG  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1243 TCCTTCACAGCATTTGGAAACTCCTGAAGACAAGTCATTTGGCCGGTCATTGACAGAGAC 1302

y    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1303 CGAAGCAAACCTGGACAGTTATCAAACAGCTTTGGAAGAGTACTCCTCGTGGCTTCTTTC 1362

y    181 TGCTGAGGACACATTGCAAGCACAAGGACAAGGAGATTTCTAATGATGGGAAGTGGTGAAAGA  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1363 AGCTGAGGATGCACTGCAAGCCCAAGGAGAGATTTCTAATGATGTCGAAGAAGTGAAAGA 1422

y    241 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCACAGGCCGGGTTGG  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1423 ACAATTTCATACTCATCATGAGGGATATATGATGGACTTGACATCCCATCCCATCAGGGACGGGTCGG 1482

y    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAAAATTATCAGAAGATGAAGA  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1483 TAATGTTCTCCAACTGGGAAGTCAACTGATTGGGAACAGGGAAATTATCAGAAGATGAAGA 1542

y    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGCCTCAGGGTAGC  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1543 AACCGAAGTGCAGGAACAAAATGAATCTCCTCAATTCAAGATGGGAATGCCTCAGGGTAGC 1602

y    421 TAGCATGGAAAAACAAAGCAATTACATTACAGAGTTTAATGATCTCCAGAATC-GAAACT  479
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
b   1603 TAGCATGGAAAAACAAAGCAATTACATTACAGAGTTCTAATGATCTCCAGAATCAGCAACT 1662

y    480 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAACAAGGAAAAATGAGAGGAGA  539
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1663 GAAAGAGTTAAATGACTGGCTAACCAAAACAGAGAGAACAAGAACAAGGAGGAAGGA 1722

y    540 GCCTCTTGGACTTGAAGACTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  599
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1723 GCCCCTTGGACCTGATATTGAAGACTTAAACGCCAAGTACAACAACATAAGGTGCTTCA 1782

y    600 AGAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCCACTCCACTCACATGGTGGTGGTAGT  659
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1783 AGAGAGACTTAGAACAAGAACAAGTCAGGGTCAATTCCTCATATGTGGTGGTAGT 1842

y    660 TGATGAATCTAGTGAGGATCACGCAACTGCTTGGGAAGAACAACTTAAGGTATTGGG  719
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1843 CGATGAATCTAGTGAGGATCACGCAACTGCTTGGGAAGAACAACTTAAGGTACTGGG 1902

y    720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTCTTTACAAGACAT  779
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1903 AGATCGATGGGCAAAACATCTGTAGGTGGACAGAGAACCGCTGGGTTCTTTCTTTACAAGACAT 1962

y    780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGCAGTGCCTTTTGTCATGGCTTTCAGA  839
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1963 CCTCCTAAAATGGCAGCGTGTTTACTGAAGAACAGCAGTGCCTTTTTGTCATGGCTTTCGGA 2022

y    840 AAAAGAAGATGCAGTGAACAGAGATTCACACACAACACAAACAACAAAATCAAAATGTT  899
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2023 GAAGGAAGATGCAGTGAACAGAGATTCACAAGTGAACACAAACAACAAAAGCAAAATGTGT 2082

y    900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAGAAAAGAGAGCAATCCAT  959
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2083 ATCAAAATCTTCAGAAAACTGGCGTGTCTTAAAAACATCTGAAAACAGAGAGCAAACCAT 2142

y    960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAAACACTGAAGAATAAGTCAGTGAC 1019
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2143 GGACAAACTCTGCTCACTCAACCAAGACCTTCCTTTCAGCGTGCTCAAGATGGGAATGCCTCAGGGTAGC 2202
```

RESULT 14
E30219

LOCUS   E30219   4402 bp   DNA   linear   PAT 18-JUN-2001
DEFINITION Shortened dystrophin.
ACCESSION   E30219
VERSION   E30219.1   GI:13017026
KEYWORDS   JP 199918467-A/2.
SOURCE   unidentified
  ORGANISM unidentified
            unclassified.
REFERENCE   1 (bases 1 to 4402)
            sinichi,T.
  TITLE   Shortened dystrophin
  JOURNAL   Patent: JP 199918467-A 2 24-NOV-1999;
            SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY
COMMENT   OS   Unidentified
            PN   JP 199918467-A/2
            PD   24-NOV-1999
            PF   08-MAY-1998 JP 1998142134
            PR
            PI   SINICHI TAKEDA
            PC   C12N15/09,A61K48/00,C12N15/00
            CC   Strandedness: Both;
            CC   Topology: Linear;
            FH   Key   Location/Qualifiers
            FT   source   1..4402
                                 /organism="Unidentified".
FEATURES   Location/Qualifiers
     source   1..4402
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Query Match   54.3%;   Score 869.2; DB 6; Length 4402;
Best Local Similarity 77.7%;   Pred. No. 2e-202;
Matches 1244;   Conservative   0;   Mismatches 33;   Indels 325;   Gaps   3;

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1095 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1154

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1155 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1214

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1215 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 1274

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGACAAGGAGATTTCTAATGATGGGAAGTGGTGAAAGA  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1275 TGCTGAGGACACATTGCAAGCACAAGGACAAGGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1334

Qy    241 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCCGGGTTGG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1335 CCAGTTTCATACTCATGATGATGGGTACATGATGATTTGACAGCCCATCAGGCCCGGGTTGG 1394

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAAAATTATCAGAAGATGAAGA  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1395 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAAAATTATCAGAAGATGAAGA 1454

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   1500  CATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGA 1559
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2271  CATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGA 2330

Qy   1560  CCATCCCAAAATGACAGAGCTCTACCAGTCTTAGCTGACCT 1601
           |||||||||||||||||||||||||||||||||||||||||
Db   2331  CCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT 2372
```

RESULT 15
```
MUSDYS
LOCUS       MUSDYS                  3275 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse dystrophin mRNA, partial cds.
ACCESSION   M18025
VERSION     M18025.1  GI:192971
KEYWORDS    dystrophin.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3275)
  AUTHORS   Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
  TITLE     Conservation of the Duchenne muscular dystrophy gene in mice and
            humans
  JOURNAL   Science 238 (4825), 347-350 (1987)
  MEDLINE   88018015
   PUBMED   3659917
COMMENT     Original source text: Mouse adult heart, cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1..3275
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS             <1..>3275
                     /note="dystrophin"
                     /codon_start=3
                     /protein_id="AAA37530.1"
                     /db_xref="GI:192972"
                     /translation="SYAFTQAAYVATSDSTQSPYPSQHLEAPRDKSLDSSLMETEVNL
                     DSYQTALEEVLSWLLSAEDTLRAQGEISNDVEEVKEQFHAHEGFMMDLTSHQGLVGNV
                     LQLGSQLVGKGKLSEDEEAEVQEQMNLLNSRWECLRVASMEKQSKLHKVLMDLQNQKL
                     KELHDWLTKTERTKRMEEEPFGPLEDLKCVQQHKVLQEDLEQEVRVNSLTHMVV
                     VVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWIVLQDILLKWQHFTEEQCLFST
                     WLSEKEDAMKNIQTSGFKDQNEMMSSLHKISTLKIDLEKKKPTMEKLSSLNQDLLSAL
                     KNKSVTQKMEIWMENFAQRWDNLTQRLEKSSAQISQAVTTTQPFLQTTVMETVTMVT
                     TREQIMVKHAQEELPPPPQKKRQITVDSELRKRLVDITELHSWITRSEAVLQSSEP
                     AVYRKEGNISLDQEKVNAIAREKAEKFRKLQDASRSAQALVEQMANEGVNAESIRQAS
                     EQLNSRWTEFCQLLSERVNWLEYQTNMITFYNQLQQLEQMTTARNLLKTQSTTLSEP
                     TAIKSQLKICKDEVNRLSALQPQIEQLKIQSLQLKEKGQPMFLDADFVAFTNHFNHI
                     FDGVRAKEKELQTIFDTLPPMRYQETMSSIRTWIQQSESKLSVPYLSVTEYEIMEERL
                     GKLQALQSSLKEQQNGPNYLSDTVKEMAKKAPSEICQKYLSEFEEIEGHWKKLSSQLV
                     ESCQKLEEHMNKLRKFQNHIKTLQKWMAEVDVFLKEWPALGDAEILKKQLKQCRLLV
                     GDIQTIQPSLNSVNEGGQKOKIKSEAELFASRLETELRELNTQWDHICRQVYTRKEALK
                     AGLDKTVSLQKDLSEMHEWMTQAEEYLERDFEYKTPDELQTAVEEMKRAKEEALQKE
                     TRVKLLTETVNSVIAHAPPSAQEALKKELETLTTNYQWLCTRLNGKCKTLEEVWACWH
                     ELLSYLEKANKWLNEVLKTMENVPAGPEITTEVLESLENLMHHSEENPQIRLLA
                     QTLTDGGVMDELINEELETFNSRWRELHEEAVRKQKLLEQSIQSAQEIEKSLHLIQES
                     LEF"
```

ORIGIN   744 bp upstream of PstI site.

```
Query Match             53.4%;  Score 854.2;  DB 10;  Length 3275;
Best Local Similarity   86.5%;  Pred. No. 9.5e-199;
Matches  954;  Conservative    0;  Mismatches  148;  Indels   1;  Gaps   1;

Qy     1  GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  60
          |||||||||||||||||||||||||||||||| | |||||||||||||||||||||||||
Db     2  GAGTTATGCCTTCCACACAGGCTGCTTATGTTGCCACCACCTCTGATTCCACACAGAGCCCCTA  61

Qy    61  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAGTCATTGGACAGTTCATTGATGGAGAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62  TCCTTCACAGCATTGGAAGCTCCCAGAGACAAGCAAGTCACTGACAGTTCATTGATGGAGAC 121
```

```
Db   1455  AACTGAAGTACAGAGACGAGATGAATCTCCTAAATTCAAGATGGGAATTGCCTCAGGGTAGC 1514

Db    421  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAATC-GAAACT  479
           ||||||||||||||||||| ||||||| |||||||||||||||||||||||||||  |||||
Db   1515  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAATCAGAAACT 1574

Qy    480  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAGAACAAGGAAAATGGAGAAGA  539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1575  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAGAACAAGGAAAATGGAGGAAGA 1634

Qy    540  GCCTCTTGGACCTGATCTTGAAGACTCCAAGTACAACAACATAAGGTGCTTCA  599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1635  GCCTCTTGGACCTGATCTTGAAGACTCCAAAGCCCAAGTACAACAACATAAGGTGCTTCA 1694

Qy    600  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACACTGTGGTGGTCGTAGT  659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1695  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCTCACCATGTGGTCGTAGT 1754

Qy    660  TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG  719
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   1755  TGATGAATCTAGTGGAGATCACGCAACTGCTGCTCTTTGGAAGAACAACTTAAGGAGTCAA 1814

Qy    720  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT  779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1815  TACTGAGTGGGAAA------------------------------------------ 1829

Qy    780  CCTTCTCAAATGGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGCTTTCAGA  839
           -------------------------------------------------------
Db   1829

Qy    840  AAAAGAAGATGCAGTGAACAAGATTCACCACACTGGCTTAAAGATCAAAATGAAATGTT  899
           -------------------------------------------------------
Db   1829

Qy    900  ATCAAGTCTTCAAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT  959
           -------------------------------------------------------
Db   1829

Qy    960  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAGAATAAGTCAGTGAC 1019
           |||                                                      |||
Db   1829  ---------------------------------------------------ATTGAA 1835

Qy   1020  CCAGAAGACACGGAAGCATGGCGGATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1079
           |||||||||||||||||||||| 
Db   1836  CCTGCACTCCGGTGACTGGC------------------------------------ 1855

Qy   1080  ACTTGAAAAGAGTACAGCACAGACCCCTGAAAGACTTCAGGAACTTCAAGAGAGGCCACGGA 1139
           |||||||||||||||||||                                       |
Db   1856  ------AGAGAAAAATAGATGAGAGCGTGGCCCTTGAAAGACTCCAAGAGAGGCCACGGA 1910

Qy   1140  TGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGATCTGGCAGCCCGTGGG 1199
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911  TGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGATCTGGCAGCCCGTGGG 1970

Qy   1200  CGATCCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGGA 1259
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1971  CGATCCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGGA 2030

Qy   1260  AATTGCGCCCTCGAAAGAACGTGAGCCACGTGAGCCGTATAACCTCAGACCTTCGCTGCC 1319
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031  AATTGCGCCCTCGAAAGAACGTGAGCCACGTGAGCCGTATAACCTCAGACCTTCGCTGCC 2090

Qy   1320  TTTGGGCATTCAGCTCGAGGTGCCGTCGAGGACCCGAGCCAGTCAGGCAGCCCCACAGGGA 1379
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2091  TTTGGGCATTCAGCTCTCACCGTATAACCTCGGAAGAACCTGAACACCAGATG 2150

Qy   1380  GAAGCTTCTGCGAGGTGCCGTCGAGGACCCGAGCCAGTCAGGGAGCCCCACAGGGA 1439
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2151  GAAGCTTCTGCGAGGTGCCGTCGAGGACCCGAGCCAGTCAGGGAGCCCCACAGGGA 2210

Qy   1440  CTTTGGTTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAGAGC 1499
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2211  CTTTGGTTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGTCCCTGGGAGAGAGC 2270
```

```
 121  TGAAGTAAACCTGGACCGTATCAAACAGCTTGAGAAGAAGTATTATCGTGGCTCTTTC   180
           || ||||| ||||| | ||||||| |||| ||||||||||  | |||||||||||||
 122  GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAGTACTTTCATGGCTCTTTC   181

 181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
      ||| |||||| ||||||| |||||||||||||||||| ||||||||| |||| ||||||||
 182  TGCCGAGGATCACATTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAAGA   241

 241  CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG   300
       |||||||||| ||||||||| | ||||||||||| ||||| | ||||||||| ||| |||
 242  ACAGTTTCATGCTCATGAGGGATTCATGATGGATCTGACATCTCATCAAGGACTTGTTGG   301

 301  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAGATGAAGA   360
      |||| |||||||| | |||||||| | | || |||| || ||| ||||||| ||||||||
 302  TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAAAATATCAGAAGATGAAGA   361

 361  AACTGAAGTACAAGAGCAGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
       | ||||||| ||||| || || ||| |||||||||||||||||||| | ||||||||||
 362  AGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAAGATGGGAAATGTCTCAGGGTAGC   421

 421  TAGCATGGAAAAACAAGCAATTACATAGAGTTTAATGATCTCCAGAATC-GAAACT   479
      ||||||||||||||||||| |||||  || |||| |||||||||||||| ||||| ||
 422  TAGCATGGAAAAACAAGCAAATTACACAAAGTTCTAATGATCTCCAGAATCAGAATT   481

 480  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAGACAAGGAAAATGGAGGAAGA   539
       ||| |  | || || ||| ||||||||  | ||||| | |||  |||||||  |  |||
 482  AAAAGAACTACATGACTGGTTAACAAAAACTGAAGAGAACTAAGAAAATGGGACGAAGA   541

 540  GCCTCTTGGACCTGATCTTGAAGAACGCCAAGTACAACACATAAGGTGCTTCA   599
      ||| | ||||||||||||||||||   |||||||||||  |||||||||||||
 542  GCCCTTTGGACCTGATCTTGAAGATCTAAAATGCCAAGTACAACAACATAAGGTGCTTCA   601

 600  AGAAGATCTAGAACAAGAACAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT   659
      |||||||||||||||||| ||| ||||||||  || |||||||||| |||||||| ||
 602  AGAAGATCTAGAACAAGAGCAGGTCAGGGTCAACTCACATGGTGGTGGTAGT   661

 660  TGATGAATCTAGTGAGATCACGCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG   719
      |||||||||  || | |||| |||| ||||||  ||||||||||||||||||||   |||
 662  TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTGGAAGAACAACTTAAGGTACTCGGG   721

 720  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT   779
      |||||||||||||||  |||| |||||||| ||||||||||| || |||||||||||| |
 722  AGATCGATGGGCAAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTTACAAGATAT   781

 780  CCTTCTCCAAATGGCAACGTCTTACTGAAGAACAGCAGTGCCTTTTTAGTGCATGGCTTTCAGA   839
       |||| |  ||||| || ||  ||||||||||||||||||||||||| ||  |||||||||||
 782  TCTTCTAAAATGCAGCATTTTACTGAAGAACAGCAGTGCCTTTTTAGTACACATGGCTTTCAGA   841

 840  AAAAGAAGATGCAGTGAACAAGCAGCAGAGATTCACACAAGATTGCTTTAAAGATCAAAATGAAATGTT   899
      ||||||||||||| |||| ||| |  |  || ||||||||| |||||||||||||||||||||||| |
 842  AAAAGAAGATGCAATGAAGACAATCATTCAGACAAGTGCTTTAAAGATCAAAATGAAATGAT   901

 900  ATCAAGTCTTCAAAACTGGCCGTTTTAAAAGCCGATCTAGAAAAGAAAAAAGCAATCCAT   959
       |||||||||||  ||| | ||||| ||||| | ||||||||||| ||||||||| |||||
 902  GTCAAGTCTTCACAAAATATCTACTTTAAAAATAGATCTAGAAAAAGAAAAAAGCCAACCAT   961

 960  GGGCAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC   1019
      || |||  | || ||||| || ||||||| |||| | ||||||| ||||||||||||||
 962  GGAAAAACTAAGTTCACTCAATCAAGATCTACTTTCGGCACTGAAAAAATAAGTCAGTGAC   1021

1020  CCAGAAGACGGAGCAATGGCCGTTTGGGATAACTTTAGTTCCAAAA   1079
       | ||||| ||| | ||| || ||| |||| ||||| |||| |||||
1022  TCAAAAGATGGAAATCTGGATGGAAAACTTTGCACAACGTTGGGACAATTTAACCCAAAA   1081

1080  ACTTGAAAAGAGTACAGCACAGA   1102
      ||||||||| ||| ||||||| |
1082  ACTTGAAAAAGAGTTCAGCACAAA   1104
```

Search completed: April 5, 2004, 02:40:17
Job time : 4182.93 secs

```
              GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

A nucleic - nucleic search, using sw model

Run on:  April  4, 2004, 16:52:43 ; Search time 3673.37 Seconds
                                     (without alignments)
                                     17079.796 Million cell updates/sec

Title:          US-09-845-416-12_COPY_900_3000
Perfect score:  2101
Sequence:       1 gagctatgcctacacacagg........ttctgcatgattctatccaa 2101

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27513289 seqs, 1493109276 residues

Total number of hits satisfying chosen parameters:  55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : EST:*
  1: em_estba:*
  2: em_esthum:*
  3: em_estin:*
  4: em_estmu:*
  5: em_estov:*
  6: em_estpl:*
  7: em_estro:*
  8: em_htc:*
  9: gb_est1:*
 10: gb_est2:*
 11: gb_htc:*
 12: gb_est3:*
 13: gb_est4:*
 14: gb_est5:*
 15: em_estfun:*
 16: em_estom:*
 17: em_gss_hum:*
 18: em_gss_inv:*
 19: em_gss_pln:*
 20: em_gss_vrt:*
 21: em_gss_fun:*
 22: em_gss_mam:*
 23: em_gss_mus:*
 24: em_gss_pro:*
 25: em_gss_rod:*
 26: em_gss_phg:*
 27: em_gss_vrl:*
 28: gb_gss1:*
 29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	787	37.5	3870	11	BC036103	BC036103 Homo sapi
2	753.8	35.9	5691	29	AY399453	AY399453 Homo sapi
3	629.8	30.0	3056	11	AK044536	AK044536 Mus muscu
4	601.8	28.6	5697	29	AY399455	AY399455 Mus muscu

c	No.	Score	%	Length	DB	ID	ID	Description
	5	509.2	24.2	3753	11	AK081426	AK081426	Mus muscu
	6	509	24.2	595	14	CB177816	CB177816	is21c01.x
	7	505	24.0	777	14	CD653550	CD653550	AGENCOURT
	8	505	24.0	824	9	AL556247	AL556247	AL556247
	9	503.4	24.0	801	14	CB991394	CB991394	AGENCOURT
	10	503.4	24.0	1098	13	BX365572	BX365572	BX365572
	11	494	23.5	620	13	BQ640063	BQ640063	he23g04.y
	12	494	23.5	797	14	CB960722	CB960722	AGENCOURT
	13	484.6	23.1	2874	29	AY408546	AY408546	Homo sapi
	14	483.8	23.0	508	14	CD701871	CD701871	EST18395
	15	480	22.8	728	14	CB228986	CB228986	AGENCOURT
	16	471.4	22.4	770	12	BG719710	BG719710	602690430
	17	471.2	22.4	2874	29	AY408548	AY408548	Mus muscu
	18	460.2	21.9	1047	14	CB850319	CB850319	MRA-0070
	19	455.8	21.7	591	13	BX490860	BX490860	DKFZp686F
	20	453.8	21.6	652	10	BB629984	BB629984	BB629984
	21	453.8	21.6	1298	11	AK087829	AK087829	Mus muscu
	22	453.8	21.6	2135	11	AK013510	AK013510	Mus muscu
	23	453.8	21.6	4437	11	AK036936	AK036936	Mus muscu
	24	452.6	21.5	5676	29	AY399454	AY399454	Pan trogl
	25	452.2	21.5	1384	11	AK075809	AK075809	Mus muscu
	26	449	21.4	704	10	BB610411	BB610411	BB610411
	27	439	20.9	578	13	BX485574	BX485574	DKFZp686L
	28	428.6	20.4	854	9	AI196693	AI196693	ui53e10.y
	29	423	20.1	665	13	BY742604	BY742604	BY742604
	30	421	20.0	520	29	CG606443	CG606443	OST28431G
	31	420.2	20.0	512	14	CF162938	CF162938	B0718G07-
	32	410.4	19.5	599	10	BB666688	BB666688	BB666688
	33	407.2	19.4	554	14	CB613696	CB613696	AMGNNUC:N
	34	404.6	19.3	3051	11	BC036095	BC036095	Homo sapi
	35	404.4	19.2	493	14	CA888041	CA888041	B0142C06-
	36	400.6	19.1	513	29	CG605968	CG605968	OST283399
	37	399.6	19.0	508	29	CG606947	CG606947	OST285335
	38	398.2	19.0	717	14	CB527785	CB527785	UI-M-FY0-
	39	397.4	18.9	495	14	CA894775	CA894775	B0187G06-
	40	392.2	18.7	650	13	BY714491	BY714491	BY714491
	41	388.6	18.5	488	29	CG597965	CG597965	OST261937
	42	388.6	18.5	542	29	CG594848	CG594848	OST253426
	43	388.6	18.5	644	13	BU313510	BU313510	603540290
	44	383.4	18.2	483	13	CA893902	CA893902	B0182B01-
	45	381.8	18.2	423	9	AA460476	AA460476	zx61e10.r

ALIGNMENTS

```
RESULT 1
BC036103
LOCUS       BC036103              3870 bp    mRNA    linear   HTC 19-NOV-2003
DEFINITION  Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker
            types), mRNA (cDNA clone IMAGE:5274415), with apparent retained
            intron.
ACCESSION   BC036103
VERSION     BC036103.1  GI:23271310
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3870)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
```

```
          Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
          Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
          Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
          Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
          Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
  JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE 22388257
   PUBMED 12477932
REFERENCE 2  (bases 1 to 3870)
  AUTHORS Strausberg,R.
  TITLE   Direct Submission
  JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
  REMARK  NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT  Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
          cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
          Toshiyuki and Piero Carninci (RIKEN)
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Sequencing Group at the Stanford Human Genome
          Center, Stanford University School of Medicine, Stanford, CA 94305
          Web site:      http://www-shgc.stanford.edu
          Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
          Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
          R. M.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 48 Row: f Column: 7
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Similarity but not
          identity to protein
          This clone has the following problem: retained intron.
FEATURES       Location/Qualifiers
     source    1..3870
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5274415"
               /tissue_type="Brain, hippocampus"
               /clone_lib="NIH_MGC_95"
               /lab_host="DH10B"
               /note="vector: pBluescript"
ORIGIN

Query Match    37.5%;  Score 787;  DB 11;  Length 3870;
Best Local Similarity   94.8%;  Pred. No. 1.6e-172;
Matches  814;  Conservative    0;  Mismatches  45;  Indels    0;  Gaps    0;

Qy      1 GAGCTATGCCTACACACAGGCTGCGTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 GAGCTATGCCTACACACAGGCTGCGTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1102

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1162

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 1222

Qy    181 TGTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 TGTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1282

Qy    241 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCCCGGGTTGG 300
            ||||||||||||||||||||||||||||||||||||||| |||||||| |||||||||
Db   1283 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACACCCCATCAGGGGCCGGGTTGG 1342
```

```
RESULT 2
AY399453

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1343 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA 1402

Qy    361 AACTGAAGTACAAGAGCAAGATGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1403 AACTGAAGTACAAGAGCAAGATGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1462

Qy    421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1463 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCCAGAATCAGAAACT 1522

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAAAACAGGAAAAATGAGGAGGAAGA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1523 GAAAGAGTTGAATGACTGGCTAACAAAAAAACAGGAAAAATGAGGAGGAAGA 1582

Qy    541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1583 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACAACATAAGGTGCTTCA 1642

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCCACTCCACTGGTGGTGGTAGT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1643 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCTCCACTCCACTGGTGGTGGTAGT 1702

Qy    661 TGATGAATCTAGTGGAGATCGACACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1703 TGATGAATCTAGTGGAGATCGACACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1762

Qy    721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTACAAGACAG 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   1763 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTACAAGACAT 1822

Qy    781 TTCTGACCAGTGGAAGCGTCTGCACCTTTCTCGAGAACTTCGTGGTGGTGGTACAGCT 840
             | ||||  ||||  |    |  |||  | ||| |||||   |     | |  | | |
Db   1823 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTTCAGA 1882

Qy    841 GAAAGATGATGAATTAAGC 859
            ||||| |||||| |||
Db   1883 AAAAGAAGATGCAGTGAAC 1901

LOCUS       AY399453              5691 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY399453
VERSION     AY399453.1  GI:39755442
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 5691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
```

```
ATURES          Location/Qualifiers
source          1..5691
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
gene            <1..>5691
                /locus_tag="HCM0229"

IGIN

Query Match      35.9%;  Score 753.8;  DB 29;  Length 5691;
Best Local Similarity  97.4%;  Pred. No. 9.6e-165;
Matches 780;  Conservative  0;  Mismatches 12;  Indels  9;  Gaps  1;

        1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
      294 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCACACGGAGCCCATT 353

       61 TCCCTGAGCCAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGAG 120
          ||||||||||       ||||||||||||||||||||||||||||||||||||||||||
      354 TCCCTGAAGAGC-------TCCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGAG 404

      121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      405 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 464

      181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      465 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 524

      241 CCAGTTCATATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      525 CCAGTTCATATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 584

      301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      585 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 644

      361 AACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      645 AACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 704

      421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      705 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 764

      481 GAAAGAGTTGAATGACTGCTACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      765 GAAAGAGTTGAATGACTGCTACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA 824

      541 GCCCTTGACCTGATCTTGAGACAGCCCAAGTACACCAAGCACAACATAAGGTGCTTCA 600
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
      825 GCCTCTTGACCTGATCTTGAGACCTCAAAGGTACACCAAGCACAACATAAGGTGCTTCA 884

      601 AGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTCAATTCTCTCACTCACATGTGGTGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      885 AGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTCAATTCTCTCACTCACATGTGGTGTAGT 944

      661 TGATGAATCTAGTGAGTCAAACATCTGTAGATGACAGAGACCGGTTCTTCTTTTACAAGTGGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      945 TGATGAATCTAGTGAGTCAAACATCTGTAGATGACAGAGACCGCTCGGTTCTTCTTTTACAAGTGGG 1004

      721 AGATCGATGGGCAAACATCTGTAGATGACAGAGACACGCTCGGTTCTTCTTTTACAAGACAG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1005 AGATCGATGGGCAAACATCTGTAGATGACAGAGACACCGCTCGGTTCTTCTTTTACAAGACAT 1064

      781 TTCTGACCAGTGGAAGCGTCT 801
          ||||||||||||||||||||
     1065 CCTTCTCAAATGGCAACGTCT 1085
```

DEFINITION   Mus musculus adult retina cDNA, RIKEN full-length enriched library,
             clone:A930019F21 product:dystrophin, muscular dystrophy, full
             insert sequence.
ACCESSION   AK044536
VERSION   AK044536.1 GI:26090404
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   MEDLINE   99279253
    PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
             Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   MEDLINE   20499374
    PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
             Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
             Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
             Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
             Fujiwake,S., Inoue,K., Togawa,K., Izawa,M., Ohara,E., Watahiki,M.,
             Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
             Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system-384-format
             sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   MEDLINE   20530913
    PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
             FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
             Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 3056)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
             Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
             Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
             Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
             Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
             Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
             Okazaki,Y., Saito,R., Saito,H., Sakai,C., Sakai,K., Sakazume,N.,
             Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
             Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
             Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
             Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to

SULT 3
044536   AK044536   3056 bp   mRNA   linear   HTC 20-SEP-2003
CUS

```
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/.
URL:http://fantom.gsc.riken.go.jp/.

FEATURES             Location/Qualifiers
     source          1..3056
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:A930019F21"
                     /db_xref="MGI:2409854"
                     /db_xref="taxon:10090"
                     /clone="A930019F21"
                     /tissue_type="retina"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    1..3056
                     /note="dystrophin, muscular dystrophy (MGD|MGI:94909,
                     GB|NM_007868, evidence: BLASTN, 100%, match=999)"
ORIGIN

Query Match      30.0%;  Score 629.8;  DB 11;  Length 3056;
Best Local Similarity   83.4%;  Pred. No. 7.2e-136;
Matches  715;  Conservative   0;  Mismatches  142;  Indels   0;  Gaps   0;

Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT    60
Db   944 GAGTTATGCCTTCACACCAGGCTGCTTATGTTGCCACCTCTGATTCCACACAGAGCCCCTA 1003
Qy    61 TCCTTCACAGCATTTGGAAGGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG   120
Db  1004 TCCTTCACAGCATTTGGAAGCTCCCAGAGACAAGTCACTGACAGTTCATTGATGGGAGAC  1063
Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
Db  1064 GGAAGTAAATCTGGATAGTTACCAAACTGCTGCTTTAGAAGAAGTACTTCTTCATGGCTTCTTTC 1123
Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGACAAGGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
Db  1124 TGCCGAGGATACATTGCGAGCACAAGGAGCACAAGAGATTTCAAATGATGTTGAAGAAGTGAAAGA 1183
Qy   241 CCAGTTCATACTCATGATGGGGTACATGATGGATTGACAGCGCCCATCAGGGCCCGGGTTGG   300
Db  1184 ACAGTTCATGTCATGAGGGATTCATGATGGATCTGACAATCTCATCAAGGACTTGTTGG   1243
Qy   301 TAATATTCTACAATTGGGAAGTGAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   360
Db  1244 TAATGTTCTCACAGTTAGGAAGTCAACTAGTTGGAAAATTATCAGAAGATGAAGA   1303
Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCATAAATTCCTAAATTCCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db  1304 AGCTGAAGTGCAAGAACAAAATGAATCTCATAAATTCCCTAAATTCCCTAAATTCAAGATGGGAATGTCTCAGGGTAGC 1363
Qy   421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAATGATCTCCAGAATCAGAAACT   480
Db  1364 TAGCATGGAAAAAACAAAGCAATTACACAGCAAATTACACAGAAAGACAACAGGAAATGAGGGAAATT 1423
Qy   481 GAAAGAGTTGAATGACTGGCTAACAACAAGAGAAAAGAAGAACAAGAAAATGAGGGAAGA   540
Db  1424 AAAAGAACTAGATGACTGGTTAACAAAAACTGAAGGAAGAACTAAGAAAATGGGAGGAAGA  1483
Qy   541 GCCCTCTTGGACCTTGGAGACCTGATCTTGAAGACTACAAAACGCAAGTACAACAACATAAGGTGCTTCA 600
Db  1484 GCCCTTTGGACCTTGATCTTGAAGATCTTGAAGATGCCAAGTACAACAACATAAGGTGCTTCA  1543
Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCCTCACATGGTGGTGGTAGT 660
Db  1544 AGAAGATCTAGAACAAGACAGGAGGTCAGGGTCAACTCACTCGCTCACTCCTCCACATGGTAGTAGTGT 1603
Qy   661 TGATGAATCTAGTGGAGATCGACCAACTCGCTGCTTGGAAGAACAACTTAAGGTATTGGG 720
```

```
Db  1604 TGATGAATCCAGCGGTGGATCATGCAACAGCTGCTTGGAGAAGAACAACTTAAGGTACTGGG 1663
Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTCTTTTACAAGACAG   780
Db  1664 AGATCGATGGGCAAATATCTGCAGATGGACTGAAGACCGCTGGATTGTTTTACAAGATAT 1723
Qy   781 TTCTGACCAGTGGAAGGGTCGCACCTTTCTGCGAGGACTTCGGTGTGGCTACAGCT 840
Db  1724 TCTTCTAAAATGGCACGCATTTTACTGAAGAACAGTGCCTTTTTAGTACTACATGGCTTTCAGA 1783
Qy   841 GAAAGATGATGAATTAA 857
Db  1784 AAAGAGAGATGCAATGA 1800
```

```
RESULT 4
AY399455

LOCUS       AY399455                5697 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY399455
VERSION     AY399455.1  GI:39755444
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..5697
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>5697
                     /locus_tag="HCM0229"
ORIGIN

Query Match      28.6%;  Score 601.8;  DB 29;  Length 5697;
Best Local Similarity   85.5%;  Pred. No. 2.8e-129;
Matches  685;  Conservative   0;  Mismatches  107;  Indels   9;  Gaps   1;

Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
Db   300 GAGTTATGCCTTCACACCAGGCTGCTTATGTTGCCACCTCTGATTCCACACAGAGCCCCTA  359
Qy    61 TCCTTCACAGCATTTGGAAGGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG  120
Db   360 TCCTTCACAG------GTCCCAGAGACAAGTCACTGACAGTTCATTGATGGGAGAC  410
Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
Db   411 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAAGTACTTCTTCATGGCTTCTTTC  470
Qy   181 TGCTGAGGACAATTGCAAGCACAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
```

```
     471 TGCCGAGGATACATTGCGAGCACAAGGAGCAGATTCAAATGATGTTGAAGAAGTGAAAGA 530
         |||| ||||  ||||||||||| |||||||| |||||||||| |||||||||||||||||
     241 CCAGTTCATACTTCATGAGGGGTACAATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 300
         |||||||||| |||||| |||| |||||| |||||| |||| ||| ||| ||  |||| |
     531 ACAGTTCATGCTCATGAGGGATCATGATGATCTGACATCTCATCAAGGACTTGTTGG 590

     301 TAATATTCTCACAATTGGGAAGTAAGCTGATTGGAACAACGAGATGAAGA 360
         |||| ||||||| || ||||||| |||||  |||| || ||| ||||||||
     591 TAATGTTCTCACAGTTAGGAAGTCAACTAGTTGAAAAGGAAATTATCAGAAGATGAAGA 650

     361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
         | |||||||| ||  | |||||||||||||||||||||||||||||||| ||||||||||
     651 AGCTGAAGTGGCAAGAACAAATGAATCTCCTAAATTCAAGATGGGAATGTCTCAGGGTAGC 710

     421 TAGCATGGAAAAACAAAGCAATTACATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
         ||||||||||||||||||||||||||| || |||| ||| |||||||||||||||||||| |
     711 TAGCCATGGAAAAACAAAGCAATTACACAAAGTTCTAATGGATCTCCAGAATCAGAAATT 770

     481 GAAAGAGTTGAATGACTGCTAACAAAAACAGAGAAGAACAAGGAAAAATGAGAGAAGA 540
         |||| ||  |||||||| ||| |||||||| |||||||||  | |||||| |||||||||
     771 AAAAGAAACTAGATGACTGGTTAACAAAAACTGAAGAGAACTAAGAAAAATGGAGGAAGA 830

     541 GCCTCTTGGACCTGATCTTGAGACTGCAAAACGCCAAGTACACAACAATCAAGGTGCTTCA 600
         |||  |||||||||||||||||| | ||||| |||||||||||||| || |||||||||||
     831 GCCCTTTGGACCTGATCTTGAGAGTCTAAAATGCCAAGTACACAACATAAGGTGCTTCA 890

     601 AGAGAGATCAGAACAAGAAGACAAGTCAGGGTCAATTCTCCTCACTCTCCACATGGTGGTGGTAGT 660
         |||| |||| |||||| || || ||||||||||| ||| ||||||||| ||||||| ||||||
     891 AGAGATCTAGAACAGGAGCAGGTCAGGGTCAACTCGCTCACTCACATGGTAGTAGTGGT 950

     661 TGATGAATCTAGTGGAGATCACGCAACTCGCTGTTTGGAAGAACAACTTAAGGTATTGGG 720
         |||||||||| ||  | |||||||||| | || || ||||||||||||||||||| ||||
     951 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGAAGAACAACTTAAGGTACTGGG 1010

     721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCCGTCGGTTCTTTTACAAGACAG 780
         ||||||||||||||| ||||| |||||||| ||||||||| |   |  ||||||||| ||
    1011 AGATCGATGGGCAAATATCTGCAGATGGACTGAAGACCCGCGAAGACCTGGATTGTTTTACAAGATAT 1070

     781 TTCTGACCAGTGGAAGCGTCT 801
         | || |  |||||||  ||||
    1071 TCTTCTAAAATGGCAGCATTT 1091
```

AUTHORS   Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
          Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
          Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
          Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
          Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,K., Matsuura,A., Kawai,J.,
          Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED   11076861
REFERENCE 4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
          FANTOM Consortium.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3753)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
          Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,T.,
          Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
          Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
          Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
          Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
          Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
          Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
          Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
          Takeda,T., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,Y.,
          Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
          Physical and Chemical Research (RIKEN), Laboratory for Genome
          Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
          RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
          Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
          URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
          Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site for further details.
          URL:http://genome.gsc.riken.go.jp/
          URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..3753
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="FANTOM_DB:C130016K19"
                      /db_xref="MGI:2413800"
                      /db_xref="taxon:10090"
                      /clone="C130016K19"
                      /tissue_type="head"
                      /clone_lib="RIKEN full-length enriched mouse cDNA library"
                      /dev_stage="16 days embryo"
     CDS   181..3054
                      /note="unnamed protein product; dystrophin related protein
                      2 (MGD|MGI:107432, GB|U43320, evidence: BLASTN, 99%,
                      match=807)
                      putative"
                      /codon_start=1
                      /protein_id="BAC38217.1"
                      /db_xref="GI:26349155"
                      /translation="MQPLVMQGCPYTLPRCHEWHAADRFHHSSLRNTCPQPQVRAAV
                      TIPAPPWDGAGDPCLSPKLLNGTVGATGPLEPSAMNLCWNEIKKKSHNLRARLEAFSD

RESULT 5
K081426
LOCUS   AK081426   3753 bp   mRNA   linear   HTC 20-SEP-2003
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
          library, clone:C130016K19 product:dystrophin related protein 2,
          full insert sequence.
ACCESSION   AK081426
VERSION   AK081426.1 GI:26349154
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE   High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED   10349636
REFERENCE 2
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
          Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED   11042159
REFERENCE 3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,

```
LSGKLQLPLREIIDWLSQKDEELSAQLPLQGDVALVQQEKETHAAFMEEVKSKGPYIS
SVLESAQAFISQHFBEELESQSESKDITSPRORIQNLSRFVWKQATVASELWEKLTAR
CVDQRHIEHTLEHILLEIGAMELSSTITQAEGVRATWEPIGDLFIDSLPEHIQAIK
LFKEEFSPVKDGVKLVNDLAHQLAISDVHLSMENSRALEQINIRWKQLQVSVAERLKQ
LQDAHRDFGPGSOHFLSTSVQVPWERAFSPNKVPYYINHQAQTTCWDHPKMTELYQTL
ADLNNIKFSAYRTAMKLRRVQKALRLDLVTLITTALEIIFNEHDLQASEHVMDVVEVIHC
LTALYERLEEERGLIVNVPLCVDMSLNWLLNVFDSGRSGKMRALSFKTGIACLCGTEV
KEKLQYLFSQVANGSSQCQQRHLGALLHEAIQVPRQLGEVAAFGGSNVRPSVRSCRF
STGKPVIERASQFLEWVNLEPQSMVWLAVLHRVTIAEQVKHQTKCSICRQCPIKGPRYR
SLKQFNVDICQTFLTGRASKGNKLHYPIMEYTPTTSSENMRDFATTLKNKFRSKQY
FSKHPQRGYLPVQSVLESDCSETPASSMLPHADTHSRIEHFASRLAEMESQNCSFFN
DSLSPDDSIDEDQYLLRHSSPITDREPAFGQQAPCSMATESKGELEKILAHLEDENRI
LQGELRRLRKWQHEAAEARPTLVEGSAEATPDHRNEELLAEARILRQHKSRLETRMQIL
EDHNKQLESLQRLRELLLQPPSESDGNGSAGSSLASSPRQSEGSHPREKGQTTPDTE
VADDVGSKSQODVSLCLEDIMEKLRHAFPSVRSSDVTANTLLAS"
```

polyA_signal 3731..3736
/note="putative"

polyA_site 3753
/note="putative"

ORIGIN

```
Query Match 24.2%; Score 509.2; DB 11; Length 3753;
Best Local Similarity 61.1%; Pred. No. 1e-107;
Matches 823; Conservative 0; Mismatches 523; Indels 0; Gaps 0;

Qy   756  ACCGCTGGGTTCTTTTTACAAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGC   815
          ||||| |||||  |  |||||||| |||||  |||||| || |||||||
Db   455  ACCTCCGGCGCTCCGCTAGAGGGCCTTCTCCAGACCTCCAGTGAAAACTTCAGCTCCGCTCC   514

Qy   816  AGGAACTTCTGGTGTGTGGCTACAGCTGAAAGATGAATTAAGCCGGGACAGGCCACCTATTG   875
          |||| |||||  || || |||| ||||   ||  |||||
Db   515  GAGAGATTATTGACTGGCTCAGCAGCCAAAGATGAGGAGTTGTCAGCTCAGCTCCCCTTGC   574

Qy   876  GAGGCGACTTTCCAGCAGTTCAGAAGCAGAGCGATGTACATAGGGCCTTCAAGAGGGAAT   935
          ||||| |||||||| ||||  ||
Db   575  AAGGGGATGTGGCCCTGGTACAACAGGAGGAGAGGACACATGCAGCTCATGTCATGAAGAAG   634

Qy   936  TGAAAACTAAAGAACTAGAGAAAACTCTACCAGGAGAATATTCTGACAG   995
          ||||||||  ||   |  ||||||||||||
Db   635  TCAAGTCAAGGGCCCCACACATCTCCTGTGCTTGAAATCGCCCAGGCTTTTCTGTGTCCC   694

Qy   996  AGCAGCCTTGGAAGGACTAGAGAAACTCTACCAGGAGAGCTGCCTCCTGAGG   1055
          ||||||||||||||
Db   695  AGCATCATTTGAAGATTAGGAGGAGTCTCAATCTGAGAGTAAAGATACCTCCCCCACAC   754

Qy   1056 AGAGAGCCCAGAAATGTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGTCAATACTGAGT   1115
          || || || |||||   ||||  || |||||
Db   755  AGCGGATTCAAAACCTTAGTCGCTTGTATGGAAGCGCAGAGAAAATGATAGAACCCTTGAAA   814

Qy   1116 GGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAGAAAAATGATAGACCCTTGAAA   1175
          ||||| |||  || |||||||  || ||
Db   815  GGGAGAAGCTGACAGCCCGCTGTGTGAGATCAGCATCGCCACATACTCTGGAAC   874

Qy   1176 GACTCCAGGAACTTCAAGAGGCACGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGG   1235
          |||||| ||||||
Db   875  ATCTATTGGAGATCCAAGGGGCCAATGAGGAGGAATTGAGCAGTACTTTGACCCAAGCAGAGG   934

Qy   1236 TGATCAAGGGATCCTGCACGCCCGTGGGCGATCTCCATTGACTCTCTCCAAGATCACC   1295
          || ||||||||||||||  |
Db   935  GAGTCCGAGCCCAGGTGGGAGCCCATAGGAGCCCATCCTTTATTGATTCCCTCCCAGAGCATA   994

Qy   1296 TCGAGAAGTCAAGGACCACTCAGGCGCTCTCGAAAGATTGCGGCTCTGAAAGAACGTGAGCCACG   1355
          || ||||| ||||||
Db   995  TCCAAGCCATCAAGCTATTCAAGAGAGAATTCTCTCCTGTGAAAGATGGGGTGAAGTTAG   1054

Qy   1356 TCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCA   1415
          |||||| ||  ||   ||||
Db   1055 TGAATGACTCTGCCCACCAGTTGCCATCTCTGATGTCACTTGCACTTGCCAATGGAGGAATTCCA   1114

Qy   1416 GCACTCTGGAAGACCTGAACAACCAGATGAGCTTCTGCAGGTGGCCGTGGAGGACCGAG   1475
          |||| |||||| |||  ||  || ||||
Db   1115 GGGCTCTGGAACAGATCAACATCCGTTGGAAACAGCTCCAGGTGTCAGTTGCTGAGAGGC   1174

Qy   1476 TCAGGGAGCTGCATGAAGCCATGAAGTGCCAGACTTTGGTCCAGCATCCTCAGCACTTTC   1535
          ||||||| || ||||||  |||   ||||||  ||
Db   1175 TTAAGCAACTCCAGGATGCCCACCGGGACTTTGGGCCTGGCTCACAAACATTCCTCTCCA   1234

Qy   1536 CGTCTGTCCAGGGTCCTGGGAGAGAGCCATCTGCCCAAACAAAGTGCCTACTATATCA   1595
          |||| ||||| || || |||||| |||| |||||
Db   1235 CTTCTGTGTGCAAGTTCCTTGGGAAAGGAAGCATTTTCCCATAAAAGTCCCATACACATCA   1294

Qy   1596 ACCACGAGACTCAAACAACTGCTGGGACCATCCCAAATGACAGAGAGTCTACCAGTCTT   1655
          |||||| |||||| |
Db   1295 ACCACCAGGCCAGACCACATGCTGGACCATCCCAAGATGACTGAGTATACCAAACCC   1354

Qy   1656 TAGCTGACCTGAATAATGTCAGATTCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACC   1715
          || ||| || |||||| |||||
Db   1355 TAGCCGATCTGAACACACATTAAGTTCTCAGACTACCTGGTAACTTAACTACAGACTCTGCGAG   1414

Qy   1716 TGCAGAAGGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACC   1775
          || ||||||||
Db   1415 TCCAGAAGGGCCCTGCGCCCTGACTGGATCGGAGGTCATGCAATTGTCATTCACTGCTTGA   1474

Qy   1776 AGCACAACCTCAAGCAAAATGACCAGCCCATGGATCTCTGCGAGATTATTAATTGTTTGA   1835
          |||
Db   1475 AGCATGACTTGCAGGCCTTGCAGGCGTCCATGTGTGATGTGGTTGGAGTGCATTCACTGCTTGA   1534

Qy   1836 CCACTATTTATGACCGCTCTGAGCAAGAGCACACAATTTGGTCAACGTCCCTCTCGG   1895
          ||  ||   |  ||
Db   1535 CTGCCTTGTATGAACGACTGGAGGAGGAAAGAGGCATCGTCCATGTCAATGTCAACGGATCC   1594

Qy   1896 TGGAATGTGTCTGAACTGGCTGTGTGAATGTTATGATACGGGACAGCAGGGAGGATCC   1955
          ||||||||| |||||||
Db   1595 TAGACATGAGCCTCAACTGGCTCCTCAATGTTTTGATAGTGGTCGCAGTGGAAAGATGC   1654

Qy   1956 GTGTCCTGTCTTTTAAAACTGGCATCATTCCCTGTGTGAAAGCAACATTGGAAGACAAGT   2015
          || |||  ||||||||
Db   1655 GAGCATTGTCCTTTAAGACTGGCATCGCCTGCCTGTGTGGCACGGAAGTGAAAGAAAAAC   1714

Qy   2016 ACAGATACCTTTCAAGCAAGTGGCAAGTTCAACATTCAGGCAGCAATTCAGGCGCAGGCTCG   2075
          |||||| ||| |||
Db   1715 TTCAGTATCTCTTCAGCCAAGTAGCCAATGGAAGTAGCCAATCAGGGCTGATCAACGCCATCTCG   1774

Qy   2076 GCCTCCTTCTGCATGATTCTATCCAA   2101
          ||||
Db   1775 GTGCCCGCTTCATGAAGCCATTCAA   1800

RESULT 6
CB177816/c
LOCUS        CB177816  595 bp  mRNA  linear  EST  31-JAN-2003
DEFINITION   is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
             similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ;, mRNA sequence.
ACCESSION    CB177816
VERSION      CB177816.1  GI:28186206
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 595)
AUTHORS      Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
             Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
             Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
             Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
             Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
             Williams,T., Jackson,Y. and Bowers,Y.
TITLE        Endocrine Pancreas Consortium
JOURNAL      Unpublished (2000)
COMMENT      Other_ESTs: is21c01.y1
             Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
             Endocrine Pancreas Consortium
             Harvard University, Howard Hughes Medical Institute
             Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
             MA 02138
             Tel: 617-495-1812
```

Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.

```
ACCESSION    CD653550
VERSION      CD653550.1  GI:31892067
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 777)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
  COMMENT    Contact: Daniela S. Gerhard, Ph.D.
             National Cancer Institute / NIH
             Office of Cancer Genomics
             Bldg. 31 Rm10A07 Bethesda, MD 20892
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
             cDNA Library Preparation: Yulan Piao and Minoru Ko
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC c         lone distribution information
             can be found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: NDAM513  row: a      column: 23
             High quality sequence stop: 715.
FEATURES             Location/Qualifiers
     source          1..777
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30426742"
                     /tissue_type="Embryonic stem cells"
                     /cell_line="WA01"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library
                     (Long)"
                     /note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
                     This is a long-transcript enriched cDNA library (Genome
                     Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
                     cell line.  Undifferentiated human ES cell line WA01/H1
                     was obtained from WiCell Research Institute, Inc.,
                     Madison, WI, cultured according to their instructions, on
                     MEF feeders. They formed round colonies with defined edges
                     and were positive for alkaline phosphatase, SSEA-4, OCT3,
                     OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45.  They are
                     negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
                     TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days
                     after plating), the ES cells from 4 X 6cm dishes were
                     treated with 1 mg/ml collagenase, type IV
                     (Invitrogen/GIBCO) for 5-10 min and gently scraped off
                     with 5 ml pipette.  RNA was purified with TRIzol Reagent
                     from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
                     (2001). [PMID:11544199]) Double-stranded cDNAs were
                     synthesized with an Oligo(dT) primer [Invitrogen:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3'] from
                     3.4g of total RNA, treated with T4 DNA polymerase, and
                     purified by ethanol-precipitation. The cDNAs were ligated
                     to Lone-linker LL-Sal4, purified by phenol/chloroform
                     extraction, and separated from free linkers by
                     Centricon-100 column.  Then, the cDNAs were amplified by
                     long-range high fidelity PCR using Ex Taq polymerase
                     (Takara) with a primer Sal4-S for 25 cycles. The products
                     were purified by phenol/chloroform extraction and
                     Centricon-100 column. The cDNAs were digested with SalI
                     and NotI enzymes and cloned into SalI/NotI site of
                     pCMV-SPORT6 plasmid vector. The average insert size is
                     about 3.6kb."
```

(Long) Homo sapiens cDNA clone IMAGE:30426742 5', mRNA sequence.

High quality sequence stop: 448.

FEATURES   Location/Qualifiers
     source   1..595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6553129"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN

Query Match   24.2%; Score 509; DB 14; Length 595;
Best Local Similarity 100.0%; Pred. No. 7.8e-108;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
        1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      509 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 450

       61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      449 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 390

      121 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      389 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 330

      181 TGCTGAGGACACATTGCAAGGACAAGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      329 TGCTGAGGACACATTGCAAGGACACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 270

      241 CCAGTTCATACTTCATGAGGGGTACATGATGAGGGGTACAGCCCATCAGCCGCCGGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      269 CCAGTTCATACTTCATGAGGGGTACATGATGAGGGGTACAGCCCATCAGCCGCCGGGTTGG 210

      301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      209 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 150

      361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      149 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 90

      421 TAGCAGAGTTGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       89 TAGCAGAGTTGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 30

      481 GAAAGAGTTGAATGACTGGCTAACAAAAA 509
          |||||||||||||||||||||||||||||
       29 GAAAGAGTTGAATGACTGGCTAACAAAAA 1
```

```
SULT 7
653550   CD653550   777 bp   mRNA   linear   EST 18-JUN-2003
CUS      AGENCOURT_14553000 NIA Human H1 Embryonic Stem Cell cDNA Library
FINITION
```

ORIGIN

Query Match   24.0%; Score 505; DB 14; Length 777;
Best Local Similarity 100.0%; Pred. No. 7.1e-107;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

```
Query Match     24.0%;  Score 505;  DB 9;  Length 824;
Best Local Similarity  100.0%;  Pred. No. 7.2e-107;
Matches  505;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy  1597 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT 1656
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT  196

Qy  1657 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1716
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   197 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT  256

Qy  1717 GCAGAAGGCCCTTGCTTGGATCTCTTGAGCTGTCAGCTGCATGTGATGCCTTGGACCA 1776
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 GCAGAAGGCCCTTGCTTGGATCTCTTGAGCTGTCAGCTGCATGTGATGCCTTGGACCA   316

Qy  1777 GCACAACCTCAAGCAAAATGACCAGCCCATGAATATCCTGCAGATTATTAATTGTTTGAC 1836
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   317 GCACAACCTCAAGCAAAATGACCAGCCCATGAATATCCTGCAGATTATTAATTGTTTGAC  376

Qy  1837 CACTATTTATGACCGGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGT 1896
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   377 CACTATTTATGACCGGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGT  436

Qy  1897 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1956
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG  496

Qy  1957 TGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA 2016
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   497 TGTCCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTA  556

Qy  2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG 2076
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   557 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG  616

Qy  2077 CCTCCTTCTGCATGATTCTATCCAA 2101
         |||||||||||||||||||||||||
Db   617 CCTCCTTCTGCATGATTCTATCCAA  641
```

RESULT 9
CB991394
LOCUS   CB991394   801 bp   mRNA   linear   EST 01-MAY-2003
DEFINITION AGENCOURT 13627932 NIH_MGC_148 Homo sapiens cDNA clone
           IMAGE:30336570 5', mRNA sequence.
ACCESSION CB991394
VERSION   CB991394.1 GI:30285818
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 801)
AUTHORS   cDNA Library Preparation: Dr. Stefan Hansson
           Tissue Procurement: Dr. Stefan Hansson
           cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
           and advice from Piero Carninci (RIKEN)
TITLE   NIH-MGC http://mgc.nci.nih.gov/,
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
COMMENT   Tissue Procurement: Dr. Stefan Hansson
           NIH-MGC http://mgc.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov

RESULT 8
AL556247
LOCUS   AL556247   824 bp   mRNA   linear   EST 31-MAY-2003
DEFINITION AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
           cDNA clone CS0DK001YB17 5-PRIME, mRNA sequence.
ACCESSION AL556247
VERSION   AL556247.2 GI:3127B051
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   On Feb 15, 2001 this sequence version replaced gi:12898746.
           Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
           Library was constructed by Life Technologies, a division of
           Invitrogen. This sequence belongs to sequence cluster 1955.r
           Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS0DK001CA09QP1.
FEATURES   Location/Qualifiers
     source   1..824
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK001YB17"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"

```
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 1098)
AUTHORS    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT    Contact: Genoscope
           Genoscope - Centre National de Sequencage
           BP 191 91006 EVRY cedex - France
           Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
           library was constructed by Life Technologies, a division of
           Invitrogen. This sequence belongs to sequence cluster 1955.r
           Contact: Peng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
           Faraday Avenue Genoscope sequence ID : CS1AK001ZB11QP1.
FEATURES             Location/Qualifiers
     source          1..1098
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK001YB21"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match       24.0%;  Score 503.4;  DB 13;  Length 1099;
Best Local Similarity   99.8%;  Pred. No. 1.8e-106;
Matches  504;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1597  CCAGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT  1656
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     204  CCAGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT  263

Qy    1657  AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT  1716
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     264  AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT  323

Qy    1717  GCAGAAGGGCCCTTTGCTTGGATCTGAGCCTGCAGCTGACTGCGCTGAGCTGATGCCTTGGACCA  1776
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     324  GCAGAAGGGCCCTTTGCTTGGKTCCTTGGKTCCTTGGTCCTGCAGCTGACTGCGCTGAGCTGATGCCTTGGACCA  383

Qy    1777  GCACAACCTCAAGCAAAATGACCAGCCCATGAGCAGAAAATGACCAGCCCATGA  1836
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     384  GCACAACCTCAAGCAAAATGACCAGCCCATGAGCAGAAAATGACCAGCCCATGA  443

Qy    1837  CACTATTATGACCGCCTGAGCAAGAGACAACAATTGGTCAACGTCCCTCTCTGCGT  1896
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     444  CACTATTATGACCGCCTGAGCAAGAGACAACAATTGGTCAACGTCCCTCTCTGCGT  503

Qy    1897  GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGAGGAGGATCCG  1956
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     504  GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGAGGAGGATCCG  563

Qy    1957  TGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAGAACAAGTA  2016
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     564  TGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAGAACAAGTA  623

Qy    2017  CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGCGCAGGCTGGG  2076
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     624  CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGCGCAGGCTGGG  683

Qy    2077  CCTCCTTCTGCATGATTCTATCCAA  2101
            |||||||||||||||||||||||||
Db     684  CCTCCTTCTGCATGATTCTATCCAA  708

RESULT 11
```

```
Db    181 AGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1848 ACCGGCTGGAGCAAGAGCACACAACAATTGGTCAACGTCCCTCTGCGTGGATATGTGTC 1907
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACCGGCTGGAGCAAGAGCACACAACAATTGGTCAACGTCCCTCTGCGTGGATATGTGTC 300
          
Qy   1908 TGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGGATCCGTGTCCTGTCTT 1967
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGGATCCGTGTCCTGTCTT 360
          
Qy   1968 TTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGCAAGTACAGATACCTTT 2027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGCAAGTACAGATACCTTT 420
          
Qy   2028 TCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCTCCTCCTGC 2087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCTCCTCCTGC 480
          
Qy   2088 ATGATTCTATCCAA 2101
          ||||||||||||||
Db    481 ATGATTCTATCCAA 494
```

```
RESULT 12
CB960722
CB960722
LOCUS       CB960722                797 bp      mRNA        linear    EST 29-APR-2003
DEFINITION  AGENCOURT_13761995 NIH_MGC_147 Homo sapiens cDNA clone
            IMAGE:30344479 5', mRNA sequence.
ACCESSION   CB960722
VERSION     CB960722.1  GI:30216839
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 797)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM380  row: n  column: 08
            High quality sequence stop: 620.
FEATURES             Location/Qualifiers
     source          1..797
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30344479"
                     /tissue_type="Human Placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_147"
                     /note="Organ: Placenta; Vector: pBluescriptR; Site_1:
                     all-XhoI; Site 2: BamH; Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for average
                     insert size 2.3 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation). Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: This is
                     a NIH_MGC library."
ORIGIN
Query Match             23.5%;  Score 494;  DB 14;  Length 797;
```

```
BQ640063
LOCUS       BQ640063                620 bp      mRNA        linear    EST 15-JUL-2002
DEFINITION  he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
            Homo sapiens cDNA clone he23g04 5', mRNA sequence.
ACCESSION   BQ640063
VERSION     BQ640063.1  GI:21764522
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 620)
  AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of human retina for the NEIBank
            Project: Retbindin, an abundant, novel retinal cDNA and alternative
            splicing of other retina-preferred gene transcripts
  JOURNAL   Mol. Vis. 8 (4), 196-204 (2002)
  MEDLINE   22103461
   PUBMED   12107411
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 23  row: g  column: 04
            Seq primer: M13Rp1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..620
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="he23g04"
                     /tissue_type="Retina"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Retina cDNA (Un-normalized,
                     unamplified): hd/he"
                     /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
                     was dissected from two 80 year old donors with no observed
                     eye disease. 100ug of total RNA was used for library
                     construction. A directionally cloned cDNA library in the
                     pSPORT1 vector (Life Technologies) was constructed at
                     Bioserve Biotechnology (Laurel MD) essentially following
                     the protocols of the SuperScript Plasmid System full
                     details of which are contained in the manufacturer's
                     instruction manual (http://www.lifetech.com/). First
                     strand synthesis was carried out using a Not I
                     primer-adapter
                     [5'-pGACTAGTTCTAGATCGGAGCGGCCGCCC(T)15-3']. EST analysis
                     was performed on the unamplified library at the NIH
                     Intramural Sequencing Center (NISC)."
ORIGIN
Query Match             23.5%;  Score 494;  DB 13;  Length 620;
Best Local Similarity   100.0%;  Pred. No. 2.5e-104;
Matches 494;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy   1608 AAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGCTGACCTGA 1667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGCTGACCTGA 60

Qy   1668 ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCC 1727
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCC 120

Qy   1728 TTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACAACCTCA 1787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACAACCTCA 180

Qy   1788 AGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATG 1847
```

Best Local Similarity 99.8%; Pred. No. 2.6e-104; Indels 1; Mismatches 0; Gaps 1;
Matches 505; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy 1597 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT 1656
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT  227

Qy 1657 AGCTGCCTGAATAATGTCAGATTCCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1716
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  228 AGCTGCCTGAATAATGTCAGATTCCAGCTTATAGGACTGCCATGAAACTCCGAAGACT  287

Qy 1717 GCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA 1776
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  288 GCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA  347

Qy 1777 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCGCAGATATTAATTGTTTGAC 1836
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  348 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCGCAGATATTAATTGTTTGAC  407

Qy 1837 CACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGTCAACGTCCCTCTCTGCGT 1896
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  408 CACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGT  467

Qy 1897 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1956
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  468 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG  527

Qy 1957 TGTCCGTGTCTTTTAAAACTGGCATCATTTCCCTGTGTGTAAAGCACATTTGGAAGACAAGTA 2016
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  528 TGTCCGTGTCTTTTAAAACTGGCATCATTTCCCTGTGTGTAAAGCACATTTGGAAGACAAGTA  587

Qy 2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGGCCAGGCTGGG 2076
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  588 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCCAGGCCAGGCTGGG  647

Qy 2077 -CCTCCTTCTGCATGATTCTATCCAA 2101
         ||||||||||||||||||||||||||
Db  648 CCTCCTTCTGCATGATTCTATCCAA  673
```

RESULT 13
AY408546
LOCUS   AY408546   2874 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION Homo sapiens DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY408546
VERSION   AY408546.1 GI:39764517
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2874)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 2874)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..2874

                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>2874
                     /gene="DRP2"
                     gene="HCM3260"
                     /locus_tag="HCM3260"

ORIGIN

Query Match   23.1%; Score 484.6; DB 29; Length 2874;
Best Local Similarity 58.3%; Pred. No. 5.4e-102;
Matches 784; Conservative 0; Mismatches 561; Indels 0; Gaps 0;

```
Qy  756 ACCGCTGGGTCTTTTACAAGACAGTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGC  815
         |||  |||||||| | ||||||||| ||
Db  275 ACCTCCGCGCTCGCCTAGAGGCCTTCCTCAGACCACCAGTGGAAAGCTTCAGCTCCCTTTTC  334

Qy  816 AGGAACTTCTGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTG  875
         ||   || ||  ||||||| |
Db  335 AAGAGATTATTGACTGGCTCAGCCCAAAGGATGGAGGAGTTGTCAGCTCAGCTGCCCCTAC  394

Qy  876 GAGGGCACTTTCCAGCAGTTCAGAAGCAGGACAGAACAGTGTACATAGGGCCTTCAAGAGGGAAT  935
         |||  |  |||||| ||
Db  395 AGGGGGATGTGGCCCTGGTGCAACAGAGGAGGAGACACATGCGGCCTTTATGGAAGAAG  454

Qy  936 TGAAAACTAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAG  995
         |  ||
Db  455 TCAAGTCTCGGGGCCCTACATCTATTCTGTGCTGGAGTCAGCTGAGTCTCCGTGTCCC  514

Qy  996 AGCAGCCTTTGGAAGGACTAGAGAAACTTCACAGGAGAGCCCAGAGAGTGCCTCCTGAGG 1055
         |||
Db  515 AGCACCCATTGAGGTTAGAGAGACGCCTCATTCTGAGAGCAAAGATACCTCCCCAGAAAC  574

Qy 1056 AGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTACTGAGT 1115
         ||  |  ||
Db  575 AGCGGATCCAGAATCCAGCGCGTTTGTATGGAAGGCAGCAGCGGAGCAGGTGGCCAGTGAACTGT  634

Qy 1116 GGGAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAA 1175
         ||
Db  635 GGGAGAGTTGACAGCCCGTGTGTGGACCAGCACCGTCACATTGAGCGGACTCTGGAGC  694

Qy 1176 GACTCCAGGAACTTCAAGAGAGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGG 1235
         |||||
Db  695 AGCTCTTGGAGATTCAAGGGCAATGGCAAGCAGGGAACTAAGCACTACTCTGAGCCAAGCGAGG  754

Qy 1236 TGATCAAGGGATCCTGGCAGCCCGTGGGCAATGGGCCCGATCCTCCTCATTGACTTCCTCAAGATCACC 1295
         ||||||
Db  755 GAGTCCGAGCCACTTGGGAGCCCATTGGGATCTCTTCATTGATTCACTCCCAGAGCACA  814

Qy 1296 TCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCCACG 1355
         |  ||
Db  815 TCCAGGCTATTAAGCTGTTCAAAGAAGAATTCTCCCCCATGAAAGATGGGATAAGTTGG  874

Qy 1356 TCAATGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTTCACCGTATAAACCTCA 1415
         ||  ||
Db  875 TGAATGATCTGGCCCACCAACTTGCCATTTCTGATGTGCACTTGTCAATGGAAATTCCC  934

Qy 1416 GCACTCGGAGAACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGGTCGAGGACCGGAG 1475
         ||  ||
Db  935 AGGCCCTGGAACAGATCAACGTCGGATGAGGAAACAACTACAGGGCGTCAGTTAGTGAGAGC  994

Qy 1476 TCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGACTTTCTTTCCA 1535
         ||  ||
Db  995 TTAAGCAGCTCCAGGATGCCCACCCGGGACTTTGGGCCTGGTCACGCACTTTCTCTCCT 1054

Qy 1536 CGTCTGTCCAGGGTGCCTCCAGGGAGAGCCATCTCGCCAAACAAGTGCCTACTATATCA 1595
         |||||
Db 1055 CCTCTGTCCAGTTCCCTGGGAAAGCAATTTCACCCATAAAGTTCCCTACTACACATCA 1114

Qy 1596 ACCACGAGACTCAAACACTTGCTGGGACCATCCCAATAAATGACAGAGCTCTACCAGTCTT 1655
         
Db 1115 ANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1174

Qy 1656 TAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGAC 1715
```

```
Db   1175  NNNCTGATCTGAACACATTAAGTTCTCAGCTTATCGCACTGCCATGAAACTCCGCGACAG  1234
                    |||| ||||| ||||||||||||||||||||||||||||||||||||||||||
Qy   1716  TGCAGAAGGCCCTTTGCTTGGATCTCTGAGCTCGTCAGCTGCTTGATGCCTTGGACC     1775

Qy   1235  TCCAGAAAGCCCTGCGCTTGGACCTGGTAACTTTAACCACAGCCCTGGAAATCTTCAATG  1294
                    |||||||||| |||||||| |||||||||||||||||||| |||||||||||||
Db   1735  GGATCTCCTGAGCCGTCAGCTGCATGTGACCTTGGACCAGCACAACCTCAAGCAAAA     1794

Qy   1776  AGCACAACCTCAAGCAAAATGACCAGCGCCCAGTGAACGCCATGAATATCCTGCAGATTAT  1835
                    ||||| || ||||||||||||||||||||||||||| ||||||||||| |||||
Db   136   GGATCTCTTGAGCCTGTCAGCTGCATGTGATGCGTGATGCCTTGGACCAGCACAACCTCAAGCAAA  195

Db   1295  AGCATGATCTGCAGGCCAGTGAGCACGTGGATGGATGCGTGAGGTCATTCACTGCCTGA   1354
                    |||| ||| |||||||||||||| ||| || |||||||||||||||||||||||
Qy   1795  TGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGATGACCGCCT  1854

Qy   1836  CCACTATTTATGATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGCG  1895
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   196   TGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGATGACCGCCT  255

Db   1355  CTGGCCTTATATGAACGTTTGGAGGAGGAAAGAGGCATCCTGGTTCAACGTGCCACTCTGTG  1414
                    ||||||||| | |||||||||||||||||||| |||||||||||||||||||||||
Qy   1855  GGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTGTGATATGTGTCTGAACTG     1914

Qy   1896  TGGATATGTGTCTGAACTGCTGTGGATACGGGACGAACAGGGAGGATCC           1955
                    |||||||||||||||||||||||||||||||||||| ||||||||
Db   256   GGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTGTGATATGTGTCTGAACTG     315

Db   1415  TGGACATGAGCCTCAATTGGCTCCTCAATGTTTTTGATAGTGCTGCGGAAAGATGC    1474
                    |||||| ||| ||||||||| | |||||||| || | || |||||||| ||
Qy   1915  GCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAAC  1974

Qy   1956  GTGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGT  2015
                    |||||||||||| ||||||||| |||||||||||||||||||||||||| ||||
Db   316   GCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTATAAAAC  375

Db   1475  GGGCATTGTCTTTTAAGACTGGCATTGCATGCATGCTTGTGTGTGGCACGGAAGTGAGGAAAAAC  1534
                    |||||||||||||||||||||||||||||||| ||||||||||| |||||||
Qy   1975  TGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGACACCTTTCAAGCA  2034

Qy   2016  ACAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGATTTGTGACCAGCGCCAGCCTGG  2075
                    |||||| |||| ||||||| | |||||||||||||||| |||||||||||||
Db   376   TGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGACACCTTTCAAGCA  435

Db   1535  TTCAGTACCCTTCAGCCAAGTGGCCAACTCAGGCAGCCAGTGTGACCAGCGCCACCTTG  1594
                    |||| ||||||||||| |||| ||
Qy   2035  AGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCCAGCCTGGGGCTCCCTTCTGCATGATTC  2094

Qy   2076  GCCTCCTTCTGCATGATTCTCATCCA           2100
                    |||||||||||||||| |||||
Db   1595  GTGTCCTGCTTCATGAGGCCATTCA           1619
                    ||||| |||
Db   436   AGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCCAGCCAGCCAGGCTGGGGCTCCCTTCTGCATGATTC  495

Qy   2095  TATCCAA                                                        2101
            |||||||
Db   496   TATCCAA                                                        502
```

RESULT 15
CB228986
LOCUS   CB228986   728 bp   mRNA   linear   EST 10-FEB-2003
DEFINITION AGENCOURT 1149247 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
            IMAGE:6884820 5', mRNA sequence.
ACCESSION   CB228986
VERSION   CB228986.1   GI:28280564
KEYWORDS   EST.
SOURCE   Macaca mulatta (rhesus monkey)
ORGANISM   Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1 (bases 1 to 728)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Eliot Spindel
            cDNA Library Preparation: CLONTECH
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3135 row: e column: 11
            High quality sequence stop: 583.
FEATURES   Location/Qualifiers
     source   1..728
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /db_xref="taxon:9544"
                     /clone="IMAGE:6884820"
                     /tissue_type="ovary"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_Rh_Ov1"
                     /note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;

RESULT 14
CD701871
LOCUS   CD701871   508 bp   mRNA   linear   EST 25-JUN-2003
DEFINITION EST18395 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD701871
VERSION   CD701871.1   GI:32232501
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 508)
AUTHORS   Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
TITLE   Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL   Unpublished (2003)
COMMENT   Contact: Yixin Zeng
            Cancer Center
            Sun Yat-sen University
            651 DongFeng Road East, Guangzhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.
FEATURES   Location/Qualifiers
     source   1..508
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="normal nasopharynx"
                     /clone_lib="human nasopharynx"
                     /note="ESTs generated from a normal nasopharynx cDNA
                     library from southern Chinese"
ORIGIN

Query Match   23.0%; Score 483.8; DB 14; Length 508;
Best Local Similarity   99.6%; Pred. No. 5.7e-102;
Matches 485; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy   1615  TTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGT  1674
                    ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   16    TTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTCTTTAGCTGACCTGAATAATGT  75
```

Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."

RIGIN

Query Match   22.8%;   Score 480; DB 14; Length 728;
Best Local Similarity   90.3%;   Pred. No. 4.8e-101;
Matches 513; Conservative   0; Mismatches 55; Indels   0; Gaps   0;

```
y    295  GGTTGGTAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGA  354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      1  GGTTGGTAATATTCACAATTGGGAAGTCAGCTGATTGGAACAGGGAAATTATCAGAAGA   60

y    355  TGAAGAAACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG  414
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     61  TGAAGAAACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAG  120

y    415  GGTAGCTAGCATGGAAAAAACAAAGCAATTACATAGAGTTTTAATGGATCTCCAGAATCA  474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    121  GGTAGCTAGCATGGAAAAAACAAAGCAATTACATAGAGTTTTAATGGATCTCCAGAATCA  180

y    475  GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAAATGGA  534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    181  GAAACTGAAAGAGTTGAATGACTGGCTGCTGCACAAAAACAGAAGAAAGAACAAGGAAAAATGGA  240

y    535  GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGT  594
              ||||||    |||||||||||||||||||||||||||||||||||||||
b    241  GAAAGAACCCCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGT  300

y    595  GCTTCAAGAAGATCTAGACAAGAACAAGTCAGGGTCAATTCTCACTCACATGGTGGT  654
              |||||||||||||||||||||||||||||||||||||||||||||||||||
b    301  GCTTCAAGAAGATCTAGACAAGAACAAGTCAGGGTCAATTCTCACTCACATGGTGGT  360

y    655  GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAGAGAACAACTTAAGGT  714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    361  GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAGAGAGAACAACTTAAGGT  420

y    715  ATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAAGAGACCCGCTGGGTTCTTTTACA  774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    421  ATTGGGAGATCGATGGGCAAACATCTGCAGATGGACAAGAGACCCGTGGGTTCTTTTACA  480

y    775  AGACAGTTCTGACCAGTGGAAGCCGTCTGCACCTTTCTCGCAGGAACTTCTCGGTGTGGCT  834
              ||||||            ||||||||    |||||
b    481  AGACATCCTTCTCAAATGCCAAGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCT  540

y    835  ACAGCTGAAAGATGATGAATTAAGCCGG  862

b    541  TTCAGAAAAAGAAGGATGTCGTGAACAGG  568
```

earch completed: April 5, 2004, 08:15:16
ob time : 3676.71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on: April 4, 2004, 16:58:13 ; Search time 101.891 Seconds
(without alignments)
11443.127 Million cell updates/sec

tle: US-09-845-416-12_COPY_900_3000
2101
1 gagctatgcctacacacagg........ttctgcatgatgctatccaa 2101

coring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

682709 seqs, 277475446 residues
searched:

otal number of hits satisfying chosen parameters: 1365418

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	DB	Length	ID	Description
1	1328.8	63.2	4	5952	US-09-687-875A-1	Sequence 1, Appli
2	1317.8	62.7	4	13977	US-09-484-970B-60	Sequence 60, Appl
3	1149.6	54.7	3	19307	US-08-836-022A-10	Sequence 10, Appl
4	1149.6	54.7	3	19307	US-09-427-048A-10	Sequence 10, Appl
5	547.4	26.1	4	6045	US-09-091-501B-7	Sequence 7, Appli
6	547.4	26.1	4	10320	US-09-091-501B-9	Sequence 9, Appli
7	517.8	24.6	4	3915	US-09-976-594-93	Sequence 93, Appl
8	79.4	3.8	4	200	US-09-091-501B-5	Sequence 5, Appli
9	78.6	3.7	4	200	US-09-091-501B-4	Sequence 4, Appli
10	78.6	3.7	4	200	US-09-091-501B-6	Sequence 6, Appli
11	76.6	3.6	1	7218	US-08-232-463-14	Sequence 14, Appl
12	46.2	2.2	4	505	US-09-621-976-15639	Sequence 15639, A
13	44.2	2.1	4	2574	US-09-668-313A-10	Sequence 10, Appl
14	44	2.1	4	1230025	US-09-198-452A-1	Sequence 1, Appli
15	43.4	2.0	4	1179	US-09-107-532A-1186	Sequence 1186, Ap
16	42.8	2.0	4	1690	US-09-620-312D-69	Sequence 69, Appl
17	42.8	2.0	3	7812	US-09-368-590-1	Sequence 1, Appli
18	40.4	1.9	4	832	US-09-621-976-2813	Sequence 2813, App
19	40.4	1.9	1	2223	US-08-257-073-4	Sequence 4, Appli
20	39.2	1.9	1	16995	US-08-961-527-82	Sequence 82, Appl
21	38.6	1.8	1	1995	US-09-425-069-3	Sequence 3, Appli
22	38.6	1.8	2	1995	US-08-317-844B-3	Sequence 3, Appli
23	38.4	1.8	4	7672	US-09-220-132-24	Sequence 24, Appl
24	38.2	1.8	4	428	US-09-668-313A-3	Sequence 3, Appli
25	38.2	1.8	6	1131	5180810-3	Patent No. 5180810
26	38.2	1.8	6	1784	5180810-2	Patent No. 5180810
27	38.2	1.8	4	4439	US-09-668-313A-17	Sequence 17, Appl
c 28	38	1.8	1394	4	US-09-247-155-76	Sequence 76, Appl
29	36.8	1.8	1886	6	521018J-1	Patent No. 5210183
30	36.6	1.7	1845	4	US-08-887-534A-22	Sequence 22, Appl
31	36.6	1.7	1845	4	US-09-527-431-22	Sequence 22, Appl
c 32	36.6	1.7	7075	4	US-08-956-171B-263	Sequence 263, App
33	36.2	1.7	289	3	US-09-007-005-17	Sequence 17, Appl
34	36.2	1.7	289	3	US-09-244-796-17	Sequence 17, Appl
35	36.2	1.7	2447	2	US-09-014-969-14	Sequence 14, Appl
c 36	36.2	1.7	11049	4	US-10-204-708-23	Sequence 23, Appl
37	36.2	1.7	168575	4	US-09-426-290-1	Sequence 1, Appli
c 38	36	1.7	724	4	US-08-956-171B-832	Sequence 832, App
39	36	1.7	2873	3	US-08-630-915A-193	Sequence 193, App
40	35.8	1.7	790	3	US-09-461-474-13	Sequence 13, Appl
41	35.6	1.7	4868	1	US-08-139-937-12	Sequence 12, Appl
42	35.6	1.7	4868	5	PCT-US93-11310-12	Sequence 12, Appl
43	35.6	1.7	5934	4	US-09-418-710-2	Sequence 2, Appli
44	35.6	1.7	8257	4	US-09-595-684B-30	Sequence 30, Appl
45	35.6	1.7	8789	1	US-08-328-254-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 654786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PEI
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc_feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match   63.2%;   Score 1328.8;   DB 4;   Length 5952;
Best Local Similarity   99.5%;   Pred. No. 0;
Matches 1333;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy  762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC 821
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3284 GGTCCCATTTGGAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC 3343

Qy  822 TTCTGGTGTGGCTACAGTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 881
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db 3344 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 3403

Qy  882 ACTTTCCAGCAGTTCAGAAGAGCAGAAACGATGTACATAGGGCCTTCAAGAGGAATTGAAAA 941
            ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 3404 ACTTTCCAGCAGTTCAGAAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGAATTGAAAA 3463

Qy  942 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
            |||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db 3464 CTAAAGAACCTGTAATCATGAGTAATCATGAGAACTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAG 3523

Qy 1002 CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAG 1061
            ||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 2
US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6425186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match   62.7%; Score 1317.8; DB 4; Length 13977;
Best Local Similarity   99.4%;
Matches 1333; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
                                 Pred. No. 0;

Qy   762 GGGTTCTTTACAAGACAGTTCTGACCAGTGAAGCGTCTGAAGCGTCTCGCACCTTCTGCGAGGAAC 821
          |||||||| ||||||||||||||||||||| ||||||||||||||||||
Db   8598 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTGCACCTTCTGCGAGGAAC 8657

Qy   822 TTCGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 881
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8658 TTCGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 8717

Qy   882 ACTTTCCAGCAGTTCAGAAGAACAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 941
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8718 ACTTTCCAGCAGTTCAGAAGAACAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 8777

Qy   942 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTCGTGACTGTACGAGCAGC 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8778 CTAAAGAACCTGTAATCATGAGTACTCTTGAGAATATTCTGACAGCAGCAGC 8837

Qy   1002 CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCGCCACGAGGCTGCTCCTGAGGAGAG 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8838 CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCGCCACGAGGCTGCTCCTGAGGAGAG 8897

Qy   1062 CCCAGAATGTCACTCGGCTTCTCACGGAGGTCAATACTGAGTGGGAAA 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   8898 CCCAGAATGTCACTCGGCTTCTCACGGAGGTCAATACTGAGTGGGAAA 8957

Qy   1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8958 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC 9017

Qy   1182 AGGAACTTCAAGAGCCACGAGCTGGACCTCAAGCTGGAGGTGATCA 1241
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   9018 AGGAACTTCAAGAGCCACGAGCTGGACCTCAAGCTGGAGGTGATCA 9077

Qy   1242 AGGGATCCTGGCAGCCCGTGGGGCGATCTCCTCCAAGATCACCTCGAGA 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   9078 AGGGATCCTGGCAGCCCGTGGGGCGATCTCCTCCAAGATCACCTCGAGA 9137

Qy   1302 AAGTCAAGGCACTTGCGGCTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAATG 1361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9138 AAGTCAAGGCACTTGCGGCTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAATG 9197

Qy   1362 ACCTTGCTCGCGCCAGCTTACCACTTGGCATTCAGTCTCCACCGTATAACCTCAGCACTC 1421

Db   3524 CTTTGGAAGGACTAGAGAAACTTCTACCAGGAGGCTGCCTCTGAGGAGAGAG 3583

Qy   1062 CCCAGAATGTCACTCGGCTTCTCACGGAGGTCAATACTGAGTGGGAAA 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3584 CCCAGAATGTCACTCGGCTTCTCACGGAGGTCAATACTGAGTGGGAAA 3643

Qy   1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3644 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC 3703

Qy   1182 AGGAACTTCAAGAGCTCAAGCTGCGCCAAGCTGAGGTGATCA 1241
          ||||||||||||||||||||||||||||||||||||||||||
Db   3704 AGGAACTTCAAGAGCTCAAGCTGCGCCAAGCTGAGGTGATCA 3763

Qy   1242 AGGGATCCTGGCGAGCCCGTGGGCGGATCTCCTCCAAGATCACCTCGAGA 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3764 AGGGATCCTGGCGAGCCCGTGGGCGGATCTCCTCCAAGATCACCTCGAGA 3823

Qy   1302 AAGTCAAGGCACTTGCGGCTCGAGGAGAAATGCGCCTCGAAAGAGAACGTGAGCCACGTCAATG 1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3824 AAGTCAAGGCACTTGCGGCTCGAGGAGAAATGCGCCTCGAAAGAGAACGTGAGCCACGTCAATG 3883

Qy   1362 ACCTTGCTCGCGCCAGCTTACCACTTGGGCATTCAGTCTCCACCGTATAACCTCAGCACTC 1421
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3884 ACCTTGCTCGCGCCAGCTTACCACTTGGGCATTCAGTCTCCACCGTATAACCTCAGCACTC 3943

Qy   1422 TGGAAGACCTGACACCAGATGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 1481
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3944 TGGAAGACCTGACACCAGATGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 4003

Qy   1482 AGCTGCCATGAAGCCCACAGGGACTTTGTCCAGCATCTCAGCACTTTCTTTTCCACGTCTG 1541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4004 AGTGCCATGAAGCCCACAGGGACTTTGTCCAGCATCTCAGCACTTTCTTTTCCACGTCTG 4063

Qy   1542 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCACTATATCACCACG 1601
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4064 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCACCACG 4123

Qy   1602 AGACTCAAACAACTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTCTTAGCTG 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4124 AGACTCAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTCTTAGCTG 4183

Qy   1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4184 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 4243

Qy   1722 AGGCCCTTTGCTTGGATCTTCTTGAGCCTGTCAGCTGACGCCTTGGACCAGCACA 1781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4244 AGGCCCTTTGCTTGGATCTTCTTGAGCCTGTCAGCTGACGCCTTGGACCAGCACA 4303

Qy   1782 ACCTCAAGAAATGACCAGCCCATGGATATCCTGCAGATATCCGAGTTATTGTTTGACCACTA 1841
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4304 ACCTCAAGAAATGACCAGCCCATGGATATCCTGCAGATATCCGAGTTATTGTTTGACCACTA 4363

Qy   1842 TTTATGACCGGCCTGGAGCAAGAGCAACAGCCACAACAATTGGTCAAGTCCCTCCTGCGTGGATA 1901
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4364 TTTATGACCGGCCTGGAGCAAGAGCAACAGCCACAACAATTGGTCAAGTCCCTCCTGCGTGGATA 4423

Qy   1902 TGTGTCTGAACTGCTGTCTGTGAATGTTTATGATACGGACGGAGATCCGTGTCC 1961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4424 TGTGTCTGAACTGCTGTCTGTGAATGTTTATGATACGGACGGAGATCCGTGTCC 4483

Qy   1962 TGTCTTTAAAACTGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 2021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4484 TGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 4543

Qy   2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTTGACCAGGCGGCCTCC 2081
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4544 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTTGACCAGGCGGCCTCC 4603

Qy   2082 TTCTGCATGATTCTATCCAA 2101
          ||||||||||||||||||||
Db   4604 TTCTGCATGATTCTATCCAA 4623

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/836,022A
  FILING DATE:
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/331,381
  FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
  NAME: Bak, Mary E.
  REGISTRATION NUMBER: 31,215
  REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 215-540-9200
  TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 19307 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: unknown
  MOLECULE TYPE: cDNA
US-08-836-022A-10

Query Match   54.7%; Score 1149.6; DB 3; Length 19307;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

```
Qy   762 GGGTTCTTTACAAGACAGTTCTGACCAGTGAAGGCTCTGCACCTTTCTCTGCAGGAAC 821
         ||||| |||||||||||||||||||||| || ||||||| ||||| |||||||||||||
Db  6096 GGTCCCATTTGGAAGCAAGTTCTGACCAGTGAAGGCGTTTGCATCTTCTTCTTCAGGAAC 6037

Qy   822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCCACCCATCGTGGTG 881
         ||||| ||||||||||||||||||||| ||||||| ||| ||||||||||||||||| ||
Db  6036 TTCTTGTTGGGCTACAGCTGAAAGATGATGAAAGATCATGAACTGAGCCGTCAGGCCGTGGTG 5977

Qy   882 ACTTTCCAGCAGTTCAGGAAGCAGAACGCATGTACATAGAGGAATTGAAAA 941
         ||||||||||||||||||||||| ||| ||||| ||||| |||||||| ||||
Db  5976 ATTTCCAGCAGTTCAGGAAGCAGAATGAATGATA ACATAGGGCCTTCAAGAGGGAATTGAAAA 5917

Qy   942 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACAGAGCAGC 1001
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5916 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTGAGACTATTTCTGAGAATATTTCTGACAGAGCAGC 5857

Qy  1002 CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAG 1061
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5856 CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAACTGCCTCCTGAGAAGAAGAG 5797

Qy  1062 CCCAGAATGTCACTCGGCTTCTCACGAAAGCAGCCTGAGGAGGTCAATACTGAGTGGGAAA 1121
         ||||||||||||||||||||||||||||||| |||||||||||||
Db  5796 CTCAGAATGTCACTCGGCTTCTCACTCGGCTCCTACGAAAGCAAGCAGGCTGAATGGGACA 5737

Qy  1122 AATTGAACCTGCACTCCGCTCAGCTGATTGGCAGCAGAAAAATAGATGAGACCCTTGAAAGACTCC 1181
         |||||||||||||||| ||||||||||||||||||||||||||| ||||||||||||||||
Db  5736 AATTGAACCTGCGCCTCAGCTGATTGGCAGCAGAAAAATAGATGAAGCTCTTGAAAGACTCC 5677

Qy  1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCA 1241
         |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  5676 AGGAACTTCAAGAGCTGCCATGAACTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 5617

Qy  1242 AGGGATCCTGGCAGCCCGTGGGCGGATCTCCATTGACTCTCTGCCAAGATCACCTCGAGA 1301
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  5616 AGGGATCCTGGCAGCCAGGTGGGGGATCCTGCCATTGACTCTCTGCCAGATCACCTTGAAA 5557

Qy  1302 AAGTCAAGGCACTTCGAGAGGAAAAATGCGCCTCTGAAAGAAGAACGTGAGCCACGTCAATG 1361
         |||||||||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||
Db  5556 AAGTCAAGGCACTTCGGGGAGAAATTGCACCTCTTAAAGAAGAATGTCAATCGTGTCAATG 5497

Qy  1362 ACCTTGCTGCCAGCTTACCACTTTGGCATTCAGCTCTATAACCTCAGCACTC 1421
         |||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  5496 ACCTTGCACATCAGCTGACCACACTGGGCATTCAGCTCTCACCTTATAACCTCAGCACTT 5437
```

```
      9198 ACCTTGCTCCGCCAGCTACCACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 9257
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

      1422 TGGAAGACCTGAACACCAGATGGAAGCTTTCGCAGTGGCCGTCGAGGACCGAGTCAGGC 1481
           ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
      9258 TGGAAGACCTGAACACCAGATGGAAGCTTTCGCAGTGGCCGTCGAGGACCGAGTCAGGC 9317

      1482 AGCTGCATGAAGCCCAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 1541
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
      9318 AGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 9377

      1542 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCACTATATCAACCACG 1601
           ||| |||||||||||||||||||||||||||||||||||| |||||||||||||||||
      9378 TCCAGG TCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCACTATATCAACCACG 9437

      1602 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 1661
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      9438 AGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 9497

      1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      9498 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 9557

      1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCAGTGATGCCTTGAGCCTTGACCAGCACA 1781
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      9558 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCAGTGATGCCTTGAGCCTTGACCAGCACA 9617

      1782 ACCTCAAGCAAAATGACCACGCCATGAGATATCCTGCAGATTATTAATTGTTTGACCACTA 1841
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
      9618 ACCTCAAGCAAAATGACCAGCCCATGAGATATCCTGCAGATTATTAATTGTTTGACCACTA 9677

      1842 TTTATGACCGGCCTGGAGCAAGAGACACAACATTTGGTCAACGTCCCTCTCTCGGCGGTGGATA 1901
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
      9678 TTTATGACCGGCCTGGAGCAAGAGCACACAACATTTGG CAACGTCCCTCTCTCGGCGGTGGATA 9737

      1902 TGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGAGACGAACAGGGAGGATCCGTGTCC 1961
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      9738 TGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGAGACGAACAGGGAGGATCCGTGTCC 9797

      1962 TGTCTTTAAAACTGGCATCATTTCCCTGTAAAGCACATTTGGAAGACAAGTACAGAT 2021
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      9798 TGTCTTTAAAACTGGCATCATTTCCCTGTAAAGCACATTTGGAAGACAAGTACAGAT 9857

      2022 ACCTTTTCAAGCAAGTGCCAAGTTCAACAGAGTTTGTGACCAGCGCAGGCT-GGGCCTC 2080
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
      9858 ACCTTTTCAAGCAAGTGCCAAGTTCAACAGAGTTTTCTGACCAGCGCAGCTGGGGCCTC 9917

      2081 CTTCTGCATGATGTCTATCCAA 2101
           |||||||||||||||||||||||
      9918 CTTCTGCATGATTCTATCCAA 9938
```

ESULT 3
3-08-836-022A-10/c
Sequence 10, Application US/08836022A
Patent No. 6001557
GENERAL INFORMATION:
  APPLICANT: Trustees of the University of Pennsylvania
  APPLICANT: Wilson, James M.
  APPLICANT: Fisher, Krishna J.
  APPLICANT: Chen, Shu-Jen
  APPLICANT: Weitzman, Matthew
  TITLE OF INVENTION: Improved Adenovirus Virus and
  NUMBER OF SEQUENCES: 10
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Howson and Howson
    STREET: Spring House Corporate Cntr, P O Box 457
    CITY: Spring House
    STATE: Pennsylvania
    COUNTRY: USA
    ZIP: 19477
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 54.7%; Score 1149.6; DB 3; Length 19307;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy   762 GGGTTCTTTACAGACAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC 821
         |||||||||||||| ||||||||||||||||||||||||||||| ||||||||||||||
Db 6096 GGTCCCATTGGAAGCAAGTTCTGACCAGTGGAAGCGTTTGCATCTTTCTCTTCAGGAAC 6037

Qy   822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAAGCCGGCAGCACCGTCAGGCACCATCGTGGTG 881
         |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6036 TTCTTGTTTGGCTACAGCTGAAAGATGATGAAGCTGAGCCGTCAGGCACCATCGTGGTG 5977

Qy   882 ACTTTCCAGAGTTCAGAAGCAGAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 941
         |||| |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db 5976 ATTTCCCAGAGTTCAGAAGCAGAGATGATATACAGGGCCTTCAAGAGGGAATTGAAAA 5917

Qy   942 CTAAAGAACCTGTAATCATGAGTACTCCTGAGACTGTACGAGTATTTCTGACAGAGCAGC 1001
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5916 CTAAAGAACCTGTAATCATGAGTACTCCTGAGACTGTGAGAATATTTCTGACAGAGCAGC 5857

Qy 1002 CTTTGGAAGGACTAGAGAAACTCTACAGGAAAGCAGCAGGAGGTCGCCTCCTGAGGAGAG 1061
         |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db 5856 CTTTGGAAGGACTAGAGAAACTCTACAGGAAAGCAGGGTGAAGAGGTCAACGTGAATGGGACA 5797

Qy 1062 CCCAGAAGTGTCACTCGGCTTCTACGAAAGCAGCAGGAGGTCAATACTGAGTGGGAAA 1121
         || |||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db 5796 CTCAGAAATGTCACTCGGCTCCTACGAAAGCAGCAGGAGGTCAACGTGAATGGGACA 5737

Qy 1122 AAGTCAAGGACACTTCGAGAGAAAATTGCGGCTCTGAAAGAGAACCCTGAAAGACTCCC 1181
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5736 AAGTCAAGGACACTTCGAGAGAAAATTGCGGCTCTGAAAGAGAACCCTGAAAGACTCCC 5677

Qy 1182 AGGAACTTCAAGAGGGCCACCGATGAGCTGACCTCCAAGCTGAGGTGATCA 1241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5676 AGGAACTTCAAGAGGGCCATGAGCTGACCTCCAAGCTGAGGTGAGGTGATCA 5617

Qy 1242 AGGGATCCTGCAGCCCGTGGCCGATCTCCATTGACTCCTCTCCAAGATCACCTCGAGA 1301
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5616 AGGGATCCTGCAGCCAGTGGGGATCTCCTCTCCAAGATCACCTTGAAA 5557

Qy 1302 AAGTCAAGGACACTTCGAGAGAAATTGCGGCTCTGAAAGAGAACGTGAGCCCTTGAAAGACTTG 1361
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5556 AAGTCAAGGACACTTCGGGAGAAATTGCACCTCTTAAAGAGAATGTCAATCGTGTCAATG 5497

Qy 1362 ACCTTGCTGCCAGTTACCACTTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
         |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db 5496 ACCTTGCACATCAGCTGACCACACTGGGCATTCAGCTCTCACCTTATAACCTCAGCACTT 5437

Qy 1422 TGGAAGACCTGAACACCAGATGGAAGCTTCGCAGGTGGCCGTCGAGGACCGAGTCAGGC 1481
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5436 TGGAAGATCTGAATACCAGATGGAGGCTTCTCACAGGTGGCCGTCGGAGGACCGAGTCAGAC 5377

Qy 1482 AGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTTTCCACGTCTG 1541
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5376 AGCTGCATGAAGCCCACAGGGACTTTGGTCCTCGCATCCCAGCACTTCCTTCCACTTCAG 5317

Qy 1542 TCCAGGGTCCCTGGGAGAGAGCCATCTGCCCAACAAAGTGCCCTACTATATCAACCACG 1601
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5316 TTCAGGGTCCCTGGGAGAGAGCCATCTCACCAAACAAAGTGCCCTACTATATCAACCACG 5257

Qy 1602 AGACTCAAACAACTTGCTGGGACCATCCAAATGACAGAGCTCTACCAGTCTTTAGCTG 1661
         |||||| ||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db 5256 AGACCCAAACAACCTTGTTGGGACCCACCCCCAAATGACAGAGCTCTCACCAGTCTTTAGCTG 5197

Qy 1662 ACCTGAATAATGTCAGATTCTCAGTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
         |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db 5196 ACCTGAATAATGTCAGGTTCTCCGGCTATAGGACTGCCATGAAGCTGCAGAAGGCTCCAGA 5137

Qy 1722 AGGCCCTTGCTTGGAATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTGGACCAGCACA 1781
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5136 AGGCCCTTGCTTGGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTGGACCAGCACA 5077

Qy 1782 ACCTCAAGCAAATGACCAGCCCATGGACATATCCTGCAGATTATTAATTGTTTGACCCTA 1841
         |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db 5076 ACCTCAAGCAAAATGACCAGCCCATGGACATATCCTGCAGATATTAATTAACTGTTTGACTACAA 5017

Qy 1842 TTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCCTCCGCGTGGATA 1901
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5016 TTTATGATCGTCTGGAGCAAGAGCACAACAATTCGTCAATGTCCCTCCTCTGTGTGGATA 4957

Qy 1902 TGTGTCTGAACTGGCTGCTCTGAAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1961
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4956 TGTGTTCCAACTGGCTTCTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 4897

Qy 1962 TGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAGAT 2021
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4896 TGTCTTTTAAAACTGGCATCATTCCTGTGTGTAAAGCACACTGGAAGACAAGTACAGAT 4837

Qy 2022 ACCTTTCAAGCAAGTGCCAAGTTGTCAACACAGATTTGTGACCAGCGCAGCTGGGCCTCC 2081
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4836 ACCTTTCAAGCAAGTGCCAAGTTGTCAACTGCTTTGTGTGACCAGCGCGTAGGCTGGGTCTTC 4777

Qy 2082 TTCTGCATGATTCTATTCCAA 2101
         ||||||||||||||||||||||
Db 4776 TTCTGCATGATTCTATTCAA 4757

RESULT 4
US-09-427-048A-10/c
Sequence 10, Application US/09427048A
Patent No. 6203975
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
Wilson, James M.
Fisher, Krishna J.
Chen, Shu-Jen
Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus and
Methods of Use Thereof
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (11)..(6037)
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: Chimeric
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (724)..(758)
;    OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match    26.1%;  Score 547.4;  DB 4;  Length 6045;
Best Local Similarity  61.4%;  Pred. No. 5.3e-160;
Matches  900;  Conservative  0;  Mismatches 556;  Indels  9;  Gaps  1;
```

```
Qy    645 ACATGGTGGTGGTAGTTGATGAATCATGAGTGAGTCACGCAACTGCTGCTTGGAAGAAC  704
              |  || || |||  || |  |     ||  | |    |  |     |  |||||| ||
Db   3270 ACAGGCAGAAGATGGTAAAAGCTTTGGGAAATTCTGAAGAGCTACTATGCTTCAACATC 3329

Qy    705 AACTTAAGGTATTGGGAGATCGGATCGATGGGCAAACATCTGTAGATGGCAGAGACCCGCTGGG  764
              |  | |     | |  | ||  |  ||  | |||  |  |    | | |||  |  |||
Db   3330 GACTGGATGATATGAACCAAAGATGGAATGACTTAAAAGCAAAATCTGCTAGCATCAGGG 3389

Qy    765 TTCTTTTACAAGACAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGAGGAACTTC  824
             || |   |  | | |  ||  | |||| || |    ||   | | |    |  ||   ||
Db   3390 CCCATTTGGAGGCCAGCGCTGAGAAGTGGAACAGGTTGCTGATGTCCTTAGAAGAACTGA 3449

Qy    825 TGGTGTGGCTACAGCTGAAAGATGAATGATGAATTAAGCCGGCACCTATTGGAGGCGACT  884
            |   || |||  | ||||| |||||| |||| |||||| |  || | |||||||  | |
Db   3450 TCAAATGGCTGAATATGAAGAGATGAAGAGCTTAAGAACAAATGCCTATTGGAGGAGATG 3509

Qy    885 TTCCAGCAGTTCAGAAGCAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTA  944
            ||||||  | ||| |||  |   | | |  | ||||| | | || |||| |  |  |
Db   3510 TTCCAGCCCTTACAGCTCCAGTATGACCATTGTAAGGCCCTGAGACGGGAGTTAAAGGAGA 3569

Qy    945 AAGAACCTGTAATCATGAGTACTTGGAAGACAAGCACATTCTGACAGAGCAGCTT 1004
            ||||| |   |  | ||||   | ||  |||  |  | ||| |  | ||||||| ||
Db   3570 AAGAATATTCTGTCCTGAATGCTGTGTCGACCAGCCCCAGTTTTCTTGGCTGATCAGCCAA 3629

Qy   1005 T--------GGAAGGACTAGAGAAACTTCACCAGGAGCTGCTCCTGAGG 1055
            |        ||| || | ||||| |  | |  | ||    |  | |||||
Db   3630 TTGAGGCCCCTGAAGAGCCAAGAGAGCAAGAGACCTACATCAAAACAGAATTAACTCCTGAGG 3689

Qy   1056 AGAGACCCAGAAATGTCACTCGGCTTCTACGAAAGCAACAGCCAGAGAAAAATAGATGAGACCCTTGAAA 1115
            ||||| |||  ||  |  |  | |  |  |  |||| |  | ||   ||||||  |  | | ||||||||||
Db   3690 AGAGAGCCCAAAAGATTGCCAAAGCCATGCGCAAAACAGTCTTCTGAAGTCAAAGAAAAAT 3749

Qy   1116 GGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGGAGAAAATAGATGGACAAGGCATTGGAGA 1175
            ||||| | | ||| |||| |||| ||||  ||||||||||||| | |  |||| |||||||||
Db   3750 GGGAAAGTCTAAATGCTAAATGCTGTAACTAGCAATGCTAGCAGCATAGAGCCACCGAGCTGAGA 3809

Qy   1176 GACTCCAGAGAACTTCAAGAGCCGGATGAGCTGGACCTCAAGCTCGCGCAAGCTGAGG 1235
            ||||| ||| ||||||||||||  |||||||||||| || || |  |  |||| ||||
Db   3810 AACTCAGAGACCTGCAGGGGAGCTATGGAGGGAATGACCTGGACGTCCACATGAAGAGGCAGAGT 3869

Qy   1236 TGATCAAGGAGGATCTGGCGCAGCCCGGGCCCGT...CCCAAGATCACC 1295
            | |||  |||  ||||||||||||||| ||||||       ||||||||
Db   3870 CCGTGCGAGAATGGCTGGAGCCCGTGGGAGGAGACTTACTCATTGACTCGCTGCAGGATCACA 3929

Qy   1296 TCGAGACAAGTCAAGGAGCACTTCGAGGAGACGAAATTGCGCCTCGAAAGAGAACGTGAGCGCACG 1355
            | ||| |  |||||||||||||  |||||| |||||||| | |   ||||| | |  | ||  |||
Db   3930 TTGAAAAAATCATGGCATTTAGAGAGGAGAAATTGCACCAATCAACTTTAAGTTAAAACGG 3989

Qy   1356 TCAATGACCTTGCTCGCGACTTACCACTTGGGCATTCAGCTTCACCGTATCAATAACCTCA 1415
            | |||||   |  |  | ||| | |  | ||| || ||||  | |  |  |||||||||| |
Db   3990 TGAATGATTTATCCAGTCAGCTGTCTCCACTTGACCTGACCGTCCATCCCTCTCTAAAGATGTCTC 4049

Qy   1416 GCACTCTGGAGAAGCTGAACCACCAGATGAAGCTTCTGCAGGTGCCGTGCGAGGACCGAG 1475
            ||  | | |||| | |||||   | ||||||| |||  ||||| || |  ||| ||| | 
Db   4050 GCCAGTTAGATGACCTTAATATGCCGATGGAAACTTTTACAGGTTCGTGGATGATCGGCC 4109
```

```
Qy   1422 TGGAAGACCTGACACCAGATGAAGCTTCTGCAGCGGCCGTCGAGGACGAGTCAGGC 1481
            ||||||| || | ||||||||| |||||| || |||  ||||| | | ||||| | |
Db   5436 TGGAAGAGATCTGAATACCAGATGAGGATGGAGGCTTCTACAGTGGCGTGTGGAGGACCGTGTCAGAC 5377

Qy   1482 AGTCGATGAAGCCCACAGGACTTGGTCAGCATCTCAGCACTTCTTTCCACGTCTG 1541
            ||||| |||||||||||||| |||||| |||||| |||||||| | ||||| |||
Db   5376 AGTCGCATGAAGCCACAGGACCACTTTGGTCCTGCATCCCAGCACTTCCTTTCCACTTCAG 5317

Qy   1542 TCCAGGGTCCTGGGAGGAGGAGCCATCCGCCAAACAAAGTGCCCTACTATATCAACCACG 1601
            | |||||||| ||||| ||  ||||||| |||| ||||||||||| || ||| | ||| |
Db   5316 TTCAGGGTCCCTGGGAGAGAGCAGCCATCCCAACAAACAAAGTGCCCTACTATCAACCACG 5257

Qy   1602 AGACTCAAACAACTTGCTCGGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGCTG 1661
            ||||  ||||  | ||| | ||||| | ||  | ||||||||||||| |||||||||||||
Db   5256 AGACCCAAACCACCACTTGTTGGGACCACCCACACCCAAATGACAGAGCTCTTACCAGTCTTTAGCTG 5197

Qy   1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAAACTGCGAAGACTGCAGA 1721
            ||||||||||||||||  ||| | |  ||||||||||||||||  | ||  |||| | ||||
Db   5196 ACCTGAATAATGTCAGTTCTCCGCGTATAGGACTGCCATGAAGCTCAGAGAGGCTCCAGA 5137

Qy   1722 AGGCCCTTGCTTGATCTCTTGAGCCTGCTGAGCCTGCATGTGATGCCTTGACCAGCACA 1781
            ||||||| |||||||||||||||||||||||||||||||||||||||| |||||||||
Db   5136 AGGCCCTTTGCTTGATCTCTTGAGCCTGCTGAGCCTGCATGTGATGCCCTGACCAGCACA 5077

Qy   1782 ACCTCAAGCAAAATGACCAGGCCCATGAGATATCCTGCAGATTATTAATTGTTGACCACTA 1841
            |||||||||||||||||||||||| |||| ||||||||||| | |||| |||||| || ||
Db   5076 ACCTCAAGCAAAATGACCAGGCCCCATGAGTATCCTGCAGATAATTAACTGTTGACTACAA 5017

Qy   1842 TTTATGACGGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCCTGCGGTGGATA 1901
            |||||||  | ||||||||||||||||||||  ||||| |||||||  |  | ||||||
Db   5016 TTTATGATCGTCTGGAGCAAGAGCACAACAACAATCTGGTCAATGTCCCTCTCTGTGTGGATA 4957

Qy   1902 TGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1961
            |||||||||||||||| | |||||||||||||||||||||||||||||||||||||||||
Db   4956 TGTGTCTCAACTGGCTTCCAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 4897

Qy   1962 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACACATTTGGAGACAAGAAGTACAGAT 2021
            ||||||||||| ||||||||||||| |||||||||||||| ||||| | |||||||||||||||
Db   4896 TGTCTTTTAAAAACTGGCATCATTTCTGTGTAAAGCACACTTGGAAGCAAGAAGTACAGAT 4837

Qy   2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGGCCTCC 2081
            |||| |||||||||||||||||||||| || |||||||||||||| | |||||||| | ||
Db   4836 ACCTTTCAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAGCCAGCGTAGGCTGGGGTCTTC 4777

Qy   2082 TTCTGCATGATTCTATCCAA 2101
            ||||||||||||||||| ||
Db   4776 TTCTGCATGATTCTATTCAA 4757
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Full length
;   OTHER INFORMATION: utrophin construct
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (724)..(758)
;   OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match        26.1%;  Score 547.4;  DB 4;  Length 10320;
Best Local Similarity 61.4%;  Pred. No. 7.7e-160;
Matches   900;  Conservative   0;  Mismatches 556;  Indels   9;  Gaps   1;

Qy    645 ACATGGTGGTGAGTGAGTGATGAATCTAGTGAGATCACGCAACTCTGCTTGGAAGAAC  704
          | | ||| |||  | | | ||| || ||| | |||| | |||  | |||||||||||||
Db   7545 ACAGGCAGAAGATGGTAAAAGCTTTGGGAAATCTCGAAGAGGCTACTATGCTTCAACATC 7604

Qy    705 AACTTAGGTATTGGGAGATCGATCGATGGCAAACATCTGTAGATGGACAGAAGACCGCTGGG  764
          ||   | | || | ||||| |  | ||  || || | | ||  |  | | |||  | |||||
Db   7605 GACTGGATGATATGAACCAAAGATGAATGACTTAAAGACAAAATCGTCTAGCATCAGGG 7664

Qy    765 TTCTTTTACAGACAGTTCTGACCAGTGGAAGCGCTGCACCTTTCTCTGCAGGAACTTC  824
          ||  || ||| ||| |  | |  || ||  | ||  | |||| | | | | | ||||||
Db   7665 CCCATTTGGAGCCAGCGCTGAGAAGTGGAACAGTTGCTGATGTCCTTAGAAGGAGACTGA 7724

Qy    825 TGGTGTGGCTACAGCTGAAAGATGAGATTAAGCCGGCAGGCACCTATTGGAGGCGACT  884
          | |||| ||| ||  ||| || | | | | || || || ||  | || | |||  ||
Db   7725 TCAAATGGCTGAATATGAAAGATGAAGAGCTTAAGCTTAAGACAAATGCTATTGGAGGAGATG 7784

Qy    885 TTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTA  944
          |||||| ||||| || || | ||  ||| |  | | |||||||| || | ||||  | ||
Db   7785 TTCCAGCCTTACAGCTCCAGTATGACCATTGTAAGCCCTGAAGTGCCTGTTAAAGGAGA 7844

Qy    945 AAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAAGTGACAGAGCAGCCTTT 1004
          || ||| | || || | |||  | | ||   | | | | || ||||| ||||||||
Db   7845 AAGAATATTCTGTCCTGAATGCTGCTGTCGCACCAGGCCCGAGTTTCTTGGCTGATCAGCCAA 7904

Qy   1005 T-------GAAGGACTAGAGAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGG 1055
          |       || ||| | |||  | | |||||| | ||||| | | | | |||||||
Db   7905 TTGAGGCCCCTGAGAGCGCAAGACCAAGAGAACCTACAATCAAAAACAGAATTAACTCCTGAGG 7964

Qy   1056 AGAGAGCCCAGAATGCACCTGGCTTCACTCGGCTTCTCACGGAGGTCAATACTGAGT 1115
          |||||||||  ||| | | | || | | | || | | | || ||| | ||| | ||
Db   7965 AGAGAGCCCAAAAAGATTGCCAAAGCCATGCGCAAAGCATGGCAAAATAGATGAGACCCTGAAA 8024

Qy   1116 GGGAAAAATTGAACCTGCACTCCGCTGACTGGCGCAGAAAAATAGATGAGACCCTTGAAA 1175
          |||||||||||||||| | | || | | || | ||| | ||| ||||| || |||||||
Db   8025 GGGAAAGTCTAAATGCTGTAACTAGCAATTGGCAAATAGCATGGACAAGGCATTGGAGA 8084

Qy   1176 GACTCCAGGGACTTCAAGAGAGGCCACCGATGAGCTCGACCTGGACGCTGGACGCTGAGG 1235
          ||| | ||||| | |||| || || || | | || | |  | |||| || ||| |||| |
Db   8085 AACTCAGAGACCTGGACGCTTATGGATGACGCTATGGACTGCTGCATGAAGGAGGCAGAGT 8144

Qy   1236 TGATCAAGGGATCCTGCAGCCCGTGGCCGTGGGCCGATCTCCTCATTGACTCTCTCCAAGATCACC 1295
          | | | ||| | | | | | || || | | || || |  |  | | || |  | || |||| ||||
Db   8145 CCGTGCGGAATGGCTGGACTGTGGGAGCTGTGGGAGACTACTACTGACTCGCTGCAGGATCACA 8204

Qy   1296 TCGAGAAGTCAAGGCACTTCGAGACGGAAATTGCGCCTTGAAAGAGAGACGTGAGCCACG 1355
          | |||||| | ||||| | | ||   ||| | | | | |||   | |  |  | || | |
Db   8205 TTGAAAAATCATGGCATTTAGAGAAGAAAATGTCACCAAATCAACTTAAAGTTAAAACGG 8264

Qy   1356 TCAATGACCTTGCTCGCCCAGCTTACCACTTTGGGACTTCCACCTTCTCTATAACCTCA 1415
          | |||| | || | | | || | | |||| |||| | | || || | | ||||| ||||
Db   8265 TGAATGATTTATCCAGTCAGTGTCTCCACTTGACCTGCATCCCTCTGAAAGATGTCTC 8324

Qy   1416 GCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCGCAGGTGCCGTCGAGGACCGAG 1475
          | | || |||| || |||| || || ||||| || || ||| | | | || ||||| |
Db   8325 GCCAGCTAGATGACTTAATATGCCATGGACGGTTTCTGTTGGATGATGCGCC 8384

Qy   1476 TCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCAATCCTCTCAGCATTTTCTTTTCCA 1535
          |||||||||| | |||||| | | || | |||||| ||||  | |||| | || || || ||||
Db   8385 TTAAACAGCTTCAGGAAGCCCACAGATTTTGGACCATCCTCTCAGCAGATTTCTCTCTA 8444
```

```
Qy   1476 TCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCTCAGCACTTTCTTTTCCA 1535
          |||| || | | | |||| | |  ||| | || | |||| || | | | | | | |||  |||
Db   4110 TTAAACAGCTTCAGGAAGCCCACAGATTTTGGACCATCCTCTCAGCAGATTTCTCTCTA 4169

Qy   1536 CGTCTGTCCAGGGTCCCTGGGAGAGAGCCACTCCTCGCCAAACAAAGTGCCCTACTATATCA 1595
          ||| ||||| || | | | | | | | | | |||| | | || || | | ||| || ||||
Db   4170 CGTCAGTCCAGCTGCCGGTGGCAAAGATCCCATTCACATAATAAAAGTGCCCTATTACATCA 4229

Qy   1596 ACCACGAGACTCAAAACAACTTGCTCGGGACCATCCCAAAATGACAGAGAGCTCTACCAGTCTT 1655
          || | || || || || || | | | || || ||| | ||| |  || | |||  | | |||
Db   4230 ACCATCAAACACAGACCAACCTTGGGACCATCCTAAAAATGACCGAACTCTTTCAATCCC 4289

Qy   1656 TAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGAC 1715
          | | | | ||| || || | | |  | | | || || | ||| | | || | | ||| | |
Db   4290 TTGCTGACCTGAATAATGTACGTTTTCTGCCTACCACCGTACAGCAATCAAAATCCGAAGAC 4349

Qy   1716 TGCAGAAGCGCCCTTGCTTGGATCTCTTGGAGCGTGTCAGCTGCATGTGATGCCTTGGACC 1775
          | ||| || | | || | || || || ||| | | | |  | | || | | ||  ||| |
Db   4350 TACAAAAAGCACTATGTTTGGATCTCTTAGAGTTGAGTACACACAAATGAAATTTCAAAC 4409

Qy   1776 AGCACAACCTCAAGCAAAATGACCAGCCATGGATCATCCTGCAGATTATTCAAATTGTTTGA 1835
          ||| || || ||| | || | | | | | || || | ||| | ||| | || ||| | || |
Db   4410 AGCACAAGTTGAACCAAAATGACCAGCCAGTCCTCAGTGTCATCAACTGTCTCA 4469

Qy   1836 CCCTATTTATGACCCGCTGAGCAAGCAGCAACAATTTGGTCACGTCCCTCCTGCG 1895
          | | || | | || | | ||| || ||| || || | | |  | |||| || || |
Db   4470 CAACAACTTATGATGGGACTTGAGCAAATGATCAAGGACCTGGTCAACGTTCACTCTGTG 4529

Qy   1896 TGGATAATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCC 1955
          | |||| ||| | | ||| || | ||| | |  | | || | ||| | ||| | ||| |||
Db   4530 TTGATATGTGTCTCAATTGGTTGGTCTGCTATGACGTGGTCGAACTGTCGGGTCGAACTA 4589

Qy   1956 GTGTCCTGTCTTTTAAACTGGCATCATTCCCTGTGTGAAAGCACATTTGGAGACAAGT 2015
          || | | | | |||| | | || |  | | |  |||  | | || | | | |||| || |
Db   4590 GAGTGCCAGAGTCTGAAGATTGGATTAATGTCTCTCTCCAAAGGTCTCTTGGAAGAAAT 4649

Qy   2016 ACAGATACCTTTTCAAGCAAGTGGCAGTTCAACAGATTTTGTGACCAGCGCGCAGGCTGG 2075
          ||| || | ||| || || ||  | | | |  | |||| | | || || | || || |||
Db   4650 ACAGATATCTCTTAAGGAGTCTGAAGATGGTCTTGGTACCAGCAGAGGCAGCTGG 4709

Qy   2076 GCCTCCTTCTGCATGATTCTATCCA 2100
          || | | | | ||| | | | ||||
Db   4710 GCCTGTGTTACTTCATGATGCCCATCCA 4734
```

```
RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
;   APPLICANT: Tinsley, Jonathon M
;   APPLICANT: Davies, Kay E
;   TITLE OF INVENTION: Utrophin gene expression
;   FILE REFERENCE: 620-42
;   CURRENT APPLICATION NUMBER: US/09/091,501B
;   CURRENT FILING DATE: 1998-06-18
;   PRIOR APPLICATION NUMBER: PCT/GB96/03156
;   PRIOR FILING DATE: 1996-12-19
;   PRIOR APPLICATION NUMBER: GB 9525962.8
;   PRIOR FILING DATE: 1995-12-19
;   PRIOR APPLICATION NUMBER: GB 9615797.9
;   PRIOR FILING DATE: 1996-07-26
;   PRIOR APPLICATION NUMBER: GB 9622174.2
;   PRIOR FILING DATE: 1996-10-24
;   NUMBER OF SEQ ID NOS: 15
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
```

```
Qy   786 ACCAGTGGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGGTGTGGCTACAGTCTGAAAG 845
             ||||||| ||| | |||||  |||||||||||| |||||||  ||||| ||||| |||||
Db   615 ACCACGTGGAAAGCTTCAGCTGCCCGTGGCCAAAGATTATTGACTGGCTCAGCCAAAAGG 674

Qy   846 ATGATGAATTAAGCCGGCAGGACCATCCCAAAATGACAGAGCTTACCAGTCTT 905
             |||||||| ||||| | | || ||  | ||||| |||  ||||||||||||||
Db   675 ATGAGGAGTTGTCAGCTCAGCCATCCTAAAATGACCGAACTCTTCAATCCC 734

Qy   906 ACGAGTGACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTA 965
             |||| || || |  |||| | || ||  || || | | | ||||||  || || |||
Db   735 AGGAGACACATGCCATGCCTTTATGTGACTCAAGGAAGTCAAGTCTCGGGGCCCTACATCTATTCTG 794

Qy   966 CTCTTGAGACTGTACCGAATATTTCTGAGCCTGTCAGCTGCATGTGATGCCTTGGACC 1025
             |||| ||| | || || | ||| |||| |  |||  |||  |  ||  ||| || |||
Db   795 TGCTGGAGTCAGCTCAGCCCTTCCTGTCCCAGCACCCATTGAGTGATTTGAGAGGCCTC 854

Qy  1026 ACCAGGAGCCCAGAGAGCTGCGTCCTCCTGAGAGAGGAGCCCAGAATGTCACTCGGCTTCTAC 1085
             || || || |  ||||||  ||  | | | || | || ||  | ||  | | |||| |||
Db   855 ATTCTGAGAGCAAAGATACCTCCCCGAAACACGGGATCCAGAATCTCAGCGCTTTGTAT 914

Qy  1086 GAAAGCAGGCTGAGAGGTCAATACTGAGTTGGTCAACGTCCACTCCGCTGACT 1145
             | | ||||| |  |||| |||||||| | |||||| | ||||||| ||||||
Db   915 GGAAGCAGGCGACGGGTGGCCAGTGAACGTTGGGAGAGTTGACAGCCCGCTGTGTGGACC 974

Qy  1146 GGCAGAGAAAAATAGATGAGACCCTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATG 1205
             | |||||| || | ||  ||  || | ||  | | ||| ||||  | |||||||||
Db   975 AGCACCGTCACATTGAGCGGACTCTTGGAGATTCAGGGG--CATGGAGG 1032

Qy  1206 AGCTGGACCTCAAGCTGCGGCCAAGTGAGGTGATCAAGGATCCTCGGCCCGTGGGGG 1265
             || ||| |||||| |  | |||| || | || || |||| |||| | |||| ||||
Db  1033 AACTAAGCACTACTCTGAGGCAAGCGAGAGTGAGGGAGTCCGAGCCCATTGGGG 1092

Qy  1266 AATCCTCATGACTCTCTGCAAGATCACCTGAGAAGTCAAGGCACTTCGAGGAGAAA 1325
             | ||||| || |||||| ||||| | ||| | | ||||| || | || ||||||
Db  1093 ATCTCTTCATTGATTCACTCACCCAGCACATCCAGCCTATTAAGCTGTTCAAAGAAGAAT 1152

Qy  1326 TTGCGCCTCTGAAAGAAACGTGAGCCACGCAATGACCTTGCTCGCCAGCTTACCCACTT 1385
             |||| |||| ||||| || ||   |  | ||||||| | | |  ||||||| |||| ||
Db  1153 TCTCCCCATGAAAGATGAGTGGAGTAAAGTTGGTGAATGATCGGGCCCAGATACTTGCCATTT 1212

Qy  1386 TGGGGCATTCAGCTCTCACCGTATAACCTGAGCACTCTGAAGACCTGAACACCAGATGGA 1445
             |||| ||||| |||  | | | |||||| ||||||  | || | | || | | |||| ||
Db  1213 CTGATGTGCACTTGTCAATGAGAATTCCCAGGCCCTGGAACAGATCAACGTCCGATGGA 1272

Qy  1446 AGCTTCTGCAGGTGGCCGTGCAGGACCAGTCAGCTGCATGAAGCCCACAGGGGACT 1505
             |||| |  |||| ||| ||| || | || || ||| |  ||||| | |||||| |||
Db  1273 AACAACTACAGGCGCAGTTGATGAGAGGCTTAAGCACTGCAGCTCCAGGATGCCACCGGGACT 1332

Qy  1506 TTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGGTCCTCGGGAGAGCCA 1565
             |||| | || ||| ||| ||||||||| | | | |||| | | | | ||||||| ||
Db  1333 TTGGGCCTGGGTCACAGCACTTTCTCTCTCTCCTCGTCTGTCCCAGGTTCCCTGGGAAGAGCAA 1392

Qy  1566 TCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGATCTCAACAACTTGCTGGGACC 1625
             | | | ||| ||| | | | | | |||| | |  |  || | | |||| | |||| |||
Db  1393 TTTCACCCAATAAAGTTCCCTACTACTACAACATCAACCAGGCTCAGACATGGCTGGGACC 1452

Qy  1626 ATCCCAAAATGACAGAGCTCTTAGCTGACGTCGAATAAATGTCAGATTCTCAG 1685
             |||||| | | | ||| | | ||| || | |||| | ||| ||| | |||||||
Db  1453 ATCCCAAGATGACAGAGTTATCACAAACCCTAGCTGATCTGAACAACATTAAGTTCTCAG 1512

Qy  1686 CTTATAGGACTGCCATGAAACTCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGA 1745
             |||| | |||| |  | |||||| |||  |  ||| | || |  ||  |||||||| ||
Db  1513 CTTATCGCACTGCCATGAAACTCCGCAGAGCTGCCAAAGCCCTGCGCGTTTGGACCTGGTAA 1572

Qy  1746 GCCTGTCGTCAGCTGCATGTGATGCCAGCACCAGCACAACCTCAAGCAAAATGACCAGCCCA 1805
             ||| |  | ||| |  | |||||| | |||  | |||||| | |||||||  | |||||| ||
Db  1573 CTTTAACCACCAGCCCTGGAAATCTTCAATGACGTGCAGGCCAGTGAGCACGTGA 1632

Qy  1806 TGGATATCTGCGAGATTATTATTGTTGACCACATTTATGACCCGTGGTAGAGAGC 1865
             ||||| | ||| || | | | ||||||| || || || ||| ||| | |||||||
Db  1633 TGGATGGTAGAGGTCATTCAGTCCTTATGAACGTTTGGAGGAGGAAA 1692

Qy  1866 ACAACAATTTGGTCAACGTCCCTCTCTCGCGTGGATATGTGTCGAACTGGCTGCTGAATG 1925
```

```
Qy  1536 CGTCTGTCCAGGGTCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCTACTATATCA 1595
Db  8445 CGTCAGTCCAGTCCGGTGGCCAAAGATCCATTTCACATAATAAAGTGCCTATTACATCA 8504

Qy  1596 ACCACGAGACTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTT 1655
Db  8505 ACCATCAAACACAGACCACCTGTTGGACCATCCTAAATGACCGAACTCTTCAATCCC 8564

Qy  1656 TAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAC 1715
Db  8565 TTGCTGACCTGAATAATGTACGTTTTTTTCTGCCTACCACGTACACGCAATCAAAATCCGAAGAC 8624

Qy  1716 TGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACC 1775
Db  8625 TACAAAAAGCACTATGTTTGGATCTCTTAGAGTTGGACAACAAATGAAATTTCAAAC 8684

Qy  1776 AGCACAACCTCAAGCAAAATGACCAGCCCATGGAATATCCTCAGATTATTCCAGATTATTAAATTGTTTGA 1835
Db  8685 AGCACAAGTTGAACCAAAATGACCAGCTCCTCCTCAGTGTTCCAGATGTCATCAACTGTCTGA 8744

Qy  1836 CCACTATTTATGACCGCCTGAGCAAGAGCAACAACAATTGGTCAACGTCCCTCTCTGCG 1895
Db  8745 CAACAACTTATGGACTGAGCAAATGTCAACGTGGTCTGTGTCAACGTTCCACTCTGTG 8804

Qy  1896 TGGATAATGTGTCGAACTGGCTGCTGCTGAATGTTTATGATACGGGACGACGAGGGGATCC 1955
Db  8805 TTGATAATGTGTCCAATTGGTGCTCCTTATGACGACGTCTATGACACTGGAAAAATTA 8864

Qy  1956 GTGTCCTCGTCTTTTAAAACTGCATCCATTCCCTGTGTAAAGCACATTGGAAGACAAGT 2015
Db  8865 GAGTGCAGAGTCGAGAGTCGAGCAAGGTCTCTTCCCAAAGTCTCTTGGAAGAAAAAT 8924

Qy  2016 ACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCGAGGCTGG 2075
Db  8925 ACAGATATCTCTTTAAGGAAGTTGCGGGGCCGCCGACAGAAAATGTGTGACCGTGACCTCG 8984

Qy  2076 GCCTCCTTCTGCATGATTCTATTCCA 2100
Db  8985 GCCTGTTACTTCATGATGCCATTCCA 9009
```

```
Query Match     24.8%;  Score 517.8; DB 4;  Length 3915;
Best Local Similarity  61.5%;  Pred. No. 7e-151;
Matches  846;  Conservative  0;  Mismatches 527;  Indels  2;  Gaps  1;

Qy   726 GATGGGCAAACATCTGAGATGGACAGAAGACCGCTGGGTTCTTTACAAGACAGTTCTG 785
             ||||||   ||||||| ||  | ||  | |  ||  ||||      | |||  |||||
Db   555 GTTGGAATGAAATGAAAAAAAAAAAAGTCCACAACCTCCGCGCCTAGAGGCCTTCTCAG 614
```

```
Db  1693 GAGGCATCCTGGTCAACGTGCCACTCTGTGTGGACATGAGCCTCAATTGCTCCTCCAATG 1752
              ||||||      |||||||||||| ||| ||||| || |||| |||||  ||| |||||
Qy  1926 TTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTAAAACTGGCATCATTT 1985

Db  1753 TTTTTGATAGTGGTCGCAGCGGAAAGATGCGGGCATTGTCTTTTAAGACTGGCATTGCAT 1812
              |||      |||| |  ||||||||| ||   ||| |||||  |||||||| ||||
Qy  1986 CCCTGTGTAAAGCACATTTGGAAGACACATTGGAAGCAAGTGGCAAGT 2045

Db  1813 GCTTGTGTGGCACGGAGTGAAGGAAGGAAAAACTCAGTACCTCTTCAGCCAAGTGGCCAACT 1872
              |    || ||||| |||||  || || ||||| ||||  || ||||| |||| | |||
Qy  2046 CAACAGGATTTGTGACCAGCGCAGGCGCAGGCGCTGGGCCTCCTTCTGCGATGATTCTATCCA 2100

Db  1873 CAGGCAGCCAGTGACCAGCGGCCACCTTGGTGTCCCGCTTCATGAGGCCATTCA 1927
              ||   |  ||  |||  | |||||||| ||||| ||||  || ||||| |||||
```

RESULT 8
US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 200
;   TYPE: DNA
;   ORGANISM: Rattus sp.
US-09-091-501B-5

Query Match   3.8%;   Score 79.4; DB 4; Length 200;
Best Local Similarity   64.3%;   Pred. No. 9.4e-15;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 565 CCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGT 624
             ||||| ||| || ||||| ||||| |||||||||| ||| | ||||
Db   16 CCTGCAAAACCTGCTTGAAGAACATAAAAGTTTGCAAAGTGACCTGCAAAGTCAGGT 75

Qy 625 CAGGGTCAATTCTCACTCACTCCTTAACTCATCATGGTGTGTGATTGTGTGTGATTGGTGG 684
             | ||||||||| |||||| |||||| || |||| |||| ||||||| ||||
Db   76 GAAGGTGAATTCCTAACTCATCATGGTGGTGATTGTGTGTGGATGAAAACAGTGGGGAGGCGC 135

Qy 685 AACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATGCGGCAAACATCGTAG 744
             || ||| || ||||||| |||| ||| ||| ||||||||| | | |||| |||
Db 136 CACAGCTGTTTTGGAAGATCAGTTACAGAAACTGGGTGAGCCGCTGAGCGCTGTATGCCG 195

Qy 745 ATGGA 749
             |||||
Db 196 CTGGA 200

RESULT 9
US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 200
;   TYPE: DNA
;   ORGANISM: Mus sp.
US-09-091-501B-4

Query Match   3.7%;   Score 78.6; DB 4; Length 200;
Best Local Similarity   62.4%;   Pred. No. 1.7e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 553 TGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCAAGAAGATCTAGA 612
             |||| ||||||| |||| |||||||| |||||||||||| || ||||||| |||
Db   4 TGACCTGCCCTCCTGCGAAGCTGCTTCAAGAACATAAAAGTTTGCAAAATGACCTTGA 63

Qy 613 ACAAGAACAAGTCAGGGTCAATTCTCACTCACTCCTTAACTCATCATGGTGTGTGAGTAGTT 672
             | ||||||| ||| ||| | ||||| || || ||||| |||||||| ||| || ||| ||
Db   64 AGCTGAACAGGTGAAGGTAAATTCCTTAACTCATCATCATGGTGTGATTGTGTGGATGAAAACAG 123

Qy 673 TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 732
             ||||||||| ||||| || || ||||||| |||| ||| ||| |||||||||| |||||
Db 124 TGGGGAGAGTGCCACAGCTCTTCTCGGAAGATCAGTCAGTTACAGAAAACTGGGTGGGTGAGCGCTGGAC 183

Qy 733 AAACATCGTAGATGGA 749
             |||||| || |||||
Db 184 AGCTGTATGCCGCTGGA 200

RESULT 10
US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 200
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match   3.7%;   Score 78.6; DB 4; Length 200;
Best Local Similarity   62.4%;   Pred. No. 1.7e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 553 TGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGA 612
             |||| |||| || |||| |||||||| |||||||||| || ||||||| |||
Db   4 TGATGTGAAATCCTACAAAAGCTGCTAGAGAGAACATAAAAGTTTGCAAAGTGATCTTGA 63
```

```
Db  1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355
          :::::::::::::::::::::::::::::::::::::::::::::::
Qy   333 GAACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAA 392
          :::::::::  ::::::::::  :::::  :::
Db  1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295
          ::::::::::::::::::::::::::::::::::::::::::::
Qy   393 ATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAG 452
          :  :::::::::::::::::::::::::::::  ::::::  :
Db  1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235
          ::::::::::::::::::::::::::::::
Qy   453 TTTTAATGGATCTCCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAG 512
          :::::  :::::   ::::::  ::::::  :::
Db  1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175
          ::::::::::::::::::::::::::::
Qy   513 AAGAAGAACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAC 572
          ::::::::::::::::::    ::
Db  1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115
          :::::::::::  ::::::
Qy   573 GCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTC 631
          :  :::  :::::  :::  :::::  |——||
Db  1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTC 1056
```

RESULT 12
US-09-621-976-15639/c
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match　　　2.2%;　Score 46.2;　DB 4;　Length 505;
Best Local Similarity　17.4%;　Pred. No. 0.00043;
Matches　54;　Conservative 131;　Mismatches 124;　Indels　2;　Gaps　1;

```
Qy   319 AAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCA 378
          ||——||    ::::  ||:|        ::::::      ||:|:|
Db   445 AATYAACRWAWGAGAWKRWKMAGSMRACARAGTTMAWGACWAMARRGWRCARGR--T 388
Qy   379 GATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAG 438
          | :|: ||::|   ||  ||::::         :::|: ::|:|
Db   387 GSWGSWKGGYRMWGRGMAAAKMRWAAAGGSYCGMTSYTSGGSKMTGRKGSWTKRKRMMTY 328
Qy   439 CAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACTCCAGAATCAGAAGAGTTGAATGACTG 498
          :::|:: |:||    ||:|:| ::::::    :::|:|
Db   327 SGMWWTSYKCTKTGKKYTGWKSKKTRWTCTSWRKYMWMGSGCWARSMKSWARSWYSMMAC 268
Qy   499 GCTAACAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGAGCCCTCTTGGACCTGATCT 558
          :::|:|     |::|::  || |:|  :::: :|::
Db   267 WCMMSASAYRARRSMYGARRSMRAGAGWRARRGKKRARGKSSMRSKMSSMRMSAGKA 208
Qy   559 TGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGA 618
          ||:|:|  :|::   ||:|:|
Db   207 RMCRRWMWSCRRMSYSCMGSKCMSCRGTCAKMWRYARYAKRYASSMGKYMMGCRWCYAKC 148
Qy   619 ACAAGTCAGGG 629
          |——||:|:|
Db   147 ARMYGYYRSRS 137
```

RESULT 13

```
Qy   613 ACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGTTGATGAATCTAG 672
          |||||||||||||||     |——||   ||||—||||||||—|||||||
Qy    64 GGCTGAACAGGTGAAGTAAATTCACTAACTCACATGGTCGTCGTTGATGAAAAACAG 123
Qy   673 TGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGTATTGGGAGATCGATGGGC 732
          ||||||| |||   |———|||    ||||||     |——|||||||
Qy   124 TGTGAGACGCGTCACAGCTATCCTAGAAGACCAGTTACAGAAACTTGGTGAGCGCTGGAC 183
Qy   733 AAACATCGTGATGGA 749
          ||||||||||||||
Qy   184 AGCAGTATGCCGTTGGA 200
```

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F.G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match　　　3.6%;　Score 76.6;　DB 1;　Length 7218;
Best Local Similarity　6.7%;　Pred. No. 8.7e-13;
Matches　28;　Conservative 236;　Mismatches 155;　Indels　0;　Gaps　0;

```
Qy   213 TTTCTAATGATGTGGAAGTGGTGAAAGAACCAGTTCATACTCATGAGGGGTACATGATGG 272
          :  :::|::    ::|:: |:||   |:|:|: |:::::||::::|:|
Db  1474 TATCTATGCAAGTAGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRR 1415
Qy   273 ATTTGACAGCCCATCAGGGCCGGGTTGGTGTAATATTCTACAATTGGGAAGTAAGCTGATTG 332
          :  :::||:|    ::|: :|||   :  ::||||    :  ::|::
```

US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
US-09-668-313A-10

Query Match   2.1%;   Score 44.2; DB 4;   Length 2574;
Best Local Similarity 54.7%;   Pred. No. 0.0056;
Matches   88;   Conservative   0; Mismatches   73; Indels   0; Gaps   0;

Qy   1553 TGGGAGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 1612
           |||||| |||||| |||||| |||||||||||
Db   1733 TGGGAGGAAGCTTACACAGCAGATGGAATCAAGTACTTCATCAACCACGTGACACAGACC 1792

Qy   1613 ACTTGCTGGGACCATCCCAAAATGCAGAGCTCTTACCAGTCTTAGCTGACCTGAATAAT 1672
           ||||||||| ||| ||| | ||||| ||| ||| |||| |
Db   1793 ACGTCCTGGATCCACCCCGTGATGAGCGCCCTGAACCTGTCCTGTGCAGAGAGAGTGAA 1852

Qy   1673 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAG 1713
           ||||||| |||| ||| || ||||
Db   1853 GAGGACTGTCCCAGAGAGCTAACAGACCCGAAAAGCTGATG 1893

RESULT 14
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)

; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t

NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

```
Query Match      2.1%;  Score 44;  DB 4;  Length 1230025;
Best Local Similarity   53.5%;  Pred. No. 0.49;
Matches   92;  Conservative   0;  Mismatches  80;  Indels   0;  Gaps   0;

Qy      342 AATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGAT 401
            |||||  |||||||||||||||||  |||| ||||| ||||||||||||||||||||||
Db   655477 AATTAGAAGAAAGAAGAGAAAGAGAAAATTGAGGATATCAAAGACTCAGATACAAAAT 656418

Qy      402 GGGAATGCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAATGG 461
            |||  ||||| |||||    |||||  ||||
Db   656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCACTAGGCAAGCAATTTATGTGG 656358

Qy      462 ATCTCCAGAATCAGAAACTGAAAGAGTTGAAATGACTGGCTAACAAAAACAGA 513
            |||||  ||| |||||||||   ||| |||||
Db   656357 CAATTAAGCAGAAAAAACTAAAAGCTTCTAAAGAGACGCGCTGGGAATAGA 656306
```

RESULT 15
US-09-107-532A-1186
; Sequence 1186, Application US/09107532A
; Patent No. 6583275
GENERAL INFORMATION:
   APPLICANT: Lynn A Doucette-Stamm and David Bush
   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
      ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
   NUMBER OF SEQUENCES: 7310
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: GENOME THERAPEUTICS CORPORATION
      STREET: 100 Beaver Street
      CITY: Waltham
      STATE: Massachusetts
      COUNTRY: USA
      ZIP: 02354
   COMPUTER READABLE FORM:
      MEDIUM TYPE: CD/ROM ISO9660
      COMPUTER: PC
      OPERATING SYSTEM: <Unknown>
      SOFTWARE: ASCII
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/107,532A
      FILING DATE: 30-Jun-1998
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 60/085,598
      FILING DATE: 14 May 1998
      APPLICATION NUMBER: 60/051571
      FILING DATE: July 2, 1997
   ATTORNEY/AGENT INFORMATION:
      NAME: Ariniello, Pamela Deneke
      REGISTRATION NUMBER: 40,489
      REFERENCE/DOCKET NUMBER: GTC-012
   TELECOMMUNICATION INFORMATION:

```
                TELEPHONE: (781)893-5007
                 TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1186:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 1179 base pairs
          TYPE: nucleic acid
          STRANDEDNESS: double
          TOPOLOGY: circular
     MOLECULE TYPE: DNA (genomic)
     HYPOTHETICAL: NO
     ANTI-SENSE: NO
     ORIGINAL SOURCE:
          ORGANISM: Enterococcus faecium
     FEATURE:
          NAME/KEY: misc feature
          LOCATION: (B) LOCATION 1...1179
     SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186
```

```
Query Match          2.1%;   Score 43.4;   DB 4;   Length 1179;
Best Local Similarity  50.7%;   Pred. No. 0.0058;
Matches  104;  Conservative   0;  Mismatches  101;  Indels   0;  Gaps   0;

Qy   471 ATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAA 530
             ||||||||||| ||||||||| ||  |  | |     ||||  |||||  |   | | ||
Db   602 AGCAAGAATAAAAGATTTGATCGATCAGACAAGAAAAATGGAGATACGATCGGAGGAA 661

Qy   531 TGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATA 590
             ||||||||||  ||  |||  |||| |  || ||||| | |  ||||||| ||||| |
Db   662 TTGTAGAAGTGCTCGTTGGAGGCGTTCCAGCTGGATTAGGAAGCTACGTACAATGGGACA 721

Qy   591 AGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGG 650
             |||||||||| ||||||    |||    ||| ||||||||||  ||||  ||||||  ||
Db   722 CGAAGCTAGATGCATGGCAAAATCGCACAAAATCGTCGTTAGTATCAATGCCTTTAAAGGCGTAG 781

Qy   651 TGGTGGTAGTTGATGAATCTAGTGG 675
             ||||||||| ||  ||||||| |||
Db   782 AATTGGGGTCGGATTCACTTCTGG 806
```

```
Search completed: April 5, 2004, 08:25:00
Job time : 105.891 secs
```

```
              GenCore version 5.1.6
    Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on on:   April  5, 2004, 04:53:30 ; Search time 509.455 Seconds
                                        (without alignments)
                                   15454.047 Million cell updates/sec

tle:         US-09-845-416-12_COPY_900_3000
rfect score: 2101
equence:     1 gagctatgctacacacagg........tctgcatgatgattctattccaa 2101

oring table: IDENTITY NUC
             Gapop 10.0 , Gapext 1.0

arched:      2470430 seqs, 1873665578 residues

tal number of hits satisfying chosen parameters:   4940860

inimum DB seq length: 0
aximum DB seq length: 2000000000

st-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries
```

Published Applications NA:*
```
 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Match	Query Length	DB Length	ID	Description
1	2101	100.0	10	3510		US-09-845-416-12	Sequence 12, Appl
2	2101	100.0	10	4476		US-09-845-416-31	Sequence 31, Appl
3	1760.4	83.8	15	5339		US-10-149-736-40	Sequence 40, Appl
4	1750.8	83.3	10	3531		US-09-845-416-10	Sequence 10, Appl
5	1750.8	83.3	10	3858		US-09-845-416-30	Sequence 30, Appl
6	1743	83.0	10	4825		US-09-845-416-9	Sequence 9, Appli
7	1743	83.0	10	4848		US-09-845-416-29	Sequence 29, Appl
8	1743	83.0	10	5060		US-09-845-416-35	Sequence 35, Appl
9	1743	83.0	10			US-09-845-416-36	Sequence 36, Appl
10	1641	78.1	10	4414		US-09-845-416-32	Sequence 32, Appl
11	1629	77.5	10	3446		US-09-845-416-14	Sequence 14, Appl
12	1629	77.5	15	5462		US-10-149-736-41	Sequence 41, Appl
13	1611.6	76.7	15	5417		US-10-149-736-39	Sequence 39, Appl
14	1602	76.2	10	3999		US-09-845-416-6	Sequence 6, Appli
15	1602	76.2	10	4966		US-09-845-416-28	Sequence 28, Appl
16	1602	76.2	10	4990		US-09-845-416-34	Sequence 34, Appl
17	1419	67.5	10	4182		US-09-845-416-2	Sequence 2, Appli
18	1419	67.5	10	5149		US-09-845-416-27	Sequence 27, Appl
19	1328.8	63.2	10	2169		US-09-845-416-4	Sequence 4, Appli
20	1328.8	63.2	15	8689		US-10-149-736-42	Sequence 42, Appl
21	1328.8	63.2	10	11058		US-09-845-416-1	Sequence 1, Appl
22	1328.8	63.2	15	11443		US-10-149-736-44	Sequence 44, Appl
23	1328.8	63.2	15	12057		US-10-149-736-47	Sequence 47, Appl
24	1328.8	63.2	9	13957		US-09-782-378A-22	Sequence 22, Appl
25	1328.8	63.2	9	13957		US-09-880-107-2284	Sequence 2284, Ap
26	1328.8	63.2	15	13957		US-10-149-736-1	Sequence 1, Appli
27	1328.8	63.2	12	14069		US-10-342-887-434	Sequence 434, App
28	1328.8	63.2	12	14082		US-10-342-887-981	Sequence 981, App
29	1328.8	63.2	15	14082		US-10-341-434-108	Sequence 108, App
30	1324	63.0	10	1821		US-09-845-416-13	Sequence 13, Appl
31	1149.6	54.7	15	13815		US-10-149-736-2	Sequence 2, Appl
32	937	44.6	10	1434		US-09-845-416-15	Sequence 15, Appl
33	785.4	37.4	15	1991		US-09-845-416-3	Sequence 3, Appli
34	777	37.0	10	1667		US-09-845-416-7	Sequence 7, Appli
35	554.2	26.4	15	11096		US-10-149-736-4	Sequence 4, Appli
36	547.4	26.1	9	10302		US-09-782-378A-23	Sequence 23, Appl
37	547.4	26.1	15	10302		US-10-149-736-3	Sequence 3, Appli
38	538.6	25.6	14	16531		US-10-101-510-667	Sequence 667, App
39	527.4	25.1	12	5106		US-10-220-120-157	Sequence 157, App
40	450	21.4	10	1340		US-09-845-416-11	Sequence 11, Appl
41	393	18.7	15	887		US-10-149-736-35	Sequence 35, Appl
42	387	18.4	15	387		US-10-149-736-32	Sequence 32, Appl
43	331	15.8	15	333		US-10-149-736-9	Sequence 9, Appli
44	327	15.6	15	327		US-10-149-736-8	Sequence 8, Appli
45	324	15.4	15	324		US-10-149-736-33	Sequence 33, Appl

ALIGNMENTS

```
RESULT 1
US-09-845-416-12
; Sequence 12, Application US/0984541612
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 12
; SOFTWARE: Patentin Ver. 2.1
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match          100.0%;   Score 2101;   DB 10;   Length 3510;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 2101;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy     1 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 959

Qy    61 TCCTTCACAGCATTTGGAAGCTCCGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTTGGAAGCTCCGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 1019

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1079

Qy   181 TGCTGAGGACGACACATTGCCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
```

Db   2160 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTGCGAAAGTCAAGGCACTTCGAGG 2219
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321 AGAAATTGCGCCTCTGAAAGAAGAACGTGAGCACTGAGCACGTCAATGACCTTGCTCGCCAGCTTAC 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2220 AGAAATTGCGCCTCTGAAAGAAGAACGTGAGCACGTCAATGACCTTGCTCGCCAGCTTAC 2279

Qy   1381 CACTTTGGGCATTCAGCTCTCCACCGTATAACCTCGAAGACCTCGAAGACCTGACCACCAG 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2280 CACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCACTCTGGAGACCTGAACACCAG 2339

Qy   1441 ATGGAAGCTTCTGCAGGTGGCCGTGCGAGGACCGAGCTGCAGGCAGCTGCATGAAGCCCACAG 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2340 ATGGAAGCTTCTGCAGGTGGCCGTGCGAGGACCGAGCTGCAGGCAGCTGCATGAAGCCCACAG 2399

Qy   1501 GGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCGTCGTCCAGGGTCCTGGGAGAG 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2400 GGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCGTCCAGGGTCCTGGGAGAG 2459

Qy   1561 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 2519

Qy   1621 GGACCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2520 GGACCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 2579

Qy   1681 CTCAGCTTATAGGACTGCCATGAAACTCCCAAGACTGCAGAAGGCCCTTGCTTGGATCT 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2580 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTGCTTGGATCT 2639

Qy   1741 CTTGAGCCTGTCAGCTGCATGGATGCCTTGCATGTCAGCTGATGCCGGTGATATGTGCTGGCTGCT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2640 CTTGAGCCTGTCAGCTGCATGGATGCCTTGCATGTCAGCTGATGCCGGTGATATGTGCTGGCTGCT 2699

Qy   1801 GCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTATTGACCGCCTGGAGCA 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700 GCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTATTGACCGCCTGGAGCA 2759

Qy   1861 AGAGCACAACAATTGGTCAACGTCCCTCTCGGTGGATATGTGTCTGAACTGGCTGCT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2760 AGAGCACAACAATTGGTCAACGTCCCTCTCGGTGGATATGTGTCTGAACTGGCTGCT 2819

Qy   1921 GAATGTTTATGATACGGGACGAGAACAGGGAGGAGATCCGTGTCCTGTCTTTAAAACTGGCAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2820 GAATGTTTATGATACGGGACGAGAACAGGGAGGAGATCCGTGTCCTGTCTTTAAAACTGGCAT 2879

Qy   1981 CATTTCCCTGTGTAAAGACATTTGGAAGACAAGTACAGATACCTTTCAAGCAAGTGGC 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2880 CATTTCCCTGTGTAAAGACATTTGGAAGACAAGTACAGATACCTTTCAAGCAAGTGGC 2939

Qy   2041 AAGTTCAACAGGATTTGTGACCAGCGCAGGCGCCTCCTTCGCATGATTCATCCA 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2940 AAGTTCAACAGGATTTGTGACCAGCGCAGGCGCCTCCTTCGCATGATTCATCCA 2999

Qy   2101 A 2101
            |
Db   3000 A 3000

Qy   1080 TGCTGAGGACACATTGCAAGCAAGAGAGAGATTTCTAATGATGTGGAAGTGTGGTGAAAGA 1139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1199

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy   421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1379

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAGAAATGGAGGAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAGAAATGGAGGAAGA 1439

Qy   541 GCCTCTTGGACCTGATCTTGAGAGACCTAAGACGTACAACATAAGGTGCTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 GCCTCTTGGACCTGATCTTGAGAGACCTAAGACGTCAACAACATAAGGTGCTTCA 1499

Qy   601 AGAGAGATTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCCACATGGTGGTGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 AGAGAGATTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCCACATGGTGGTGTAGT 1559

Qy   661 TGATGAATCTAGTGGACAAACATCTGTAGAATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1560 TGATGAATCTAGTGGACAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAG 1619

Qy   721 AGATCGATGGGACAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1620 AGATCGATGGGACAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAG 1679

Qy   781 TTCTGACCAGTGAAGGCTCTCGCACCTTTCTCGTGTTGTGGCTACAGCT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1680 TTCTGACCAGTGAAGGCTCTCGCACCTTTCTCGGTGTTGTGGCTACAGCT 1739

Qy   841 GAAAGATGATGAATTAAGCCGGCAGCCACCCATTGGAGGCGACTTTCCAGCAGTTCAGAA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1740 GAAAGATGATGAATTAAGCCGGCAGCCACCCATTGGAGGCGACTTTCCAGCAGTTCAGA 1799

Qy   901 GCGAACCATGTACATGAGGCCTTCAAGAGGGAATTGAAAACTAAAGACCTGTAATCAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1800 GCGAACCATGTACATGAGGCCTTCAAGAGGGAATTGAAAACTAAAGACCTGTAATCAT 1859

Qy   961 GAGTACTCTTGAGACTGTACGAGAATATTTCTGACAGAGGACTAGAGAA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1860 GAGTACTCTTGAGACTGTACGAGAATATTTCTGACAGAGGACTAGAGAA 1919

Qy   1021 ACTCTACCAGAGACCCCAGAGAGCTGCCTCCTGAGGAGAATGTCACTCGGCT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1920 ACTCTACCAGAGACCCCAGAGAGCTGCCTCCTGAGGAGAAATGTCACTCGGCT 1979

Qy   1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTCGGAAAAATTGAACCTGCACTCCGC 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1980 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTCGGAAAAATTGAACCTGCACTCCGC 2039

Qy   1141 TGACTGGCAGAGAAAAATAGATGAGACCCTGAAAGACTTCAGAGAACTTCAGAAGGGCCAC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2040 TGACTGGCAGAGAAAAATAGATGAGACCCTGAAAGACTTCAGAGAACTTCAGAAGGGCCAC 2099

Qy   1201 GGATGAGCTGGACCTCAAGCTGCGCCAAGCTGCGCCAAGCTGCGCCAAGGATCTGGCCCGCCCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2100 GGATGAGCTGGACCTCAAGCTGCGCCAAGCTGCGCCAAGCTGCGCCAAGGATCTGGCCCGCCCGT 2159

Qy   1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTGCGAAAGTCAAGGCACTTCGAGG 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 4476
TYPE: DNA
ORGANISM: Homo sapiens
5-09-845-416-31

Query Match          100.0%;   Score 2101;   DB 10;   Length 4476;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 2101;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
   1 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1656 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1715

  61 TCCTTCACAGCATTTGGAAGCTTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  120
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1716 TCCTTCACAGCATTTGGAAGCTTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1775

 121 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTTCTTTC  180
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1776 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTTCTTTC 1835

 181 TGCTGAGGACACAATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1836 TGCTGAGGACACAATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1895

 241 CCAGTTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG  300
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1896 CCAGTTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG 1955

 301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1956 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2015

 361 AACTGAAGTACAAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2016 AACTGAAGTACAAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2075

 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  480
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2076 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2135

 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGA  540
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2136 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGA 2195

 541 GCCTCTTGAGACCTGATCTTGAAGAACTACAACAACATAAGGTGCTTCA  600
     |||||||||||||||||||||||||||||||||||||||||||||||||
2196 GCCTCTTGAGACCTGATCTTGAAGAACTACAACAACATAAGGTGCTTCA 2255

 601 AGAAGATCTAGAACAAGAACAAGAAGTCAGGGCAACTGTCTGCTTCACTCACAGTTAGT  660
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2256 AGAAGATCTAGAACAAGAACAAGAAGTCAGGGCAACTGTCTGCTTCACTCACAGTTAGT 2315

 661 TGATGAATCTAGTGAGATCGAGAATCACCGCAACTGTCTGCTTTGGAAGAACTAAGGTATTGGG  720
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2316 TGATGAATCTAGTGAGATCGAGAATCACCGCAACTGTCTGCTTTGGAAGAACTAAGGTATTGGG 2375

 721 AGATCGATGGGCAAACATCGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAG  780
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2376 AGATCGATGGGCAAACATCGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAG 2435

 781 TTCTGACCAGTGGAAGCGTCGACCTTTCTCTGCAGAACTTCTGGTGTGGCTACAGCT  840
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2436 TTCTGACCAGTGGAAGCGTCGACCTTTCTCTGCAGAACTTCTGGTGTGGCTACAGCT 2495

 841 GAAAGATGATGAATTAAGCCGGCAGCCGACCTATTGGAAGAGCGACTTCCAGCAGTTCAGAA  900
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2496 GAAAGATGATGAATTAAGCCGGCAGCCGACCTATTGGAAGAGCGACTTCCAGCAGTTCAGAA 2555

 901 GCAGAACGATGTACATAGGGCCTTCAAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCAT  960
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2556 GCAGAACGATGTACATAGGGCCTTCAAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCAT 2615

 961 GAGTACTCTTGAGACTGTACGAGAATATTCTGACAGAGCAGCCTTTGGAAGAGGACTAGAGAA 1020
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2616 GAGTACTCTTGAGACTGTACGAGAATATTCTGACAGAGCAGCCTTTGGAAGAGGACTAGAGAA 2675

1021 ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAGAGCCCAGAATGTCACTCGGCT 1080
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2676 ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAGAGCCCAGAATGTCACTCGGCT 2735

1081 TCTACGAAGACAGGCTGGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCCACTCCGC 1140
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2736 TCTACGAAGACAGGCTGGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCCACTCCGC 2795

1141 TGACTGGCGAGAAAAATAGATGAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGCCAC 1200
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2796 TGACTGGCGAGAAAAATAGATGAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGCCAC 2855

1201 GGATGAGCTGGACCTGCCCAAGCTGCAGGTGATCAAGAGCTGAGGTCCTGCAGCCCGT 1260
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2856 GGATGAGCTGGACCTGCCCAAGCTGCAGGTGATCAAGAGCTGAGGTCCTGCAGCCCGT 2915

1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCACCTTCGAGG 1320
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2916 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCACCTTCGAGG 2975

1321 AGAAATTCGCCTTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTAC 1380
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2976 AGAAATTCGCCTTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTAC 3035

1381 CACTTTGGGCATTCAGCTCTCCACGGTATAACCTCGGAGACTCGAACACCAG 1440
     |||||||||||||||||||||||||||||||||||||||||||||||||||
3036 CACTTTGGGCATTCAGCTCTCCACGGTATAACCTCGGAGACTCGAACACCAG 3095

1441 ATGGAAGTTCTTGCAGTGGCCGTCGAGGACCGAGCTCAGGCAGTCCATGAGAGCCCACAG 1500
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3096 ATGGAAGTTCTTGCAGTGGCCGTCGAGGACCGAGCTCAGGCAGTCCATGAGAGCCCACAG 3155

1501 GGACTTTGGTCCAGCATCCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGAGAG 1560
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||
3156 GGACTTTGGTCCAGCATCCAGCACTTTCTTTCCACGTCTGTCCAGGTCCTGGAGAG 3215

1561 AGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCGAGACTCAAACACTTGCTG 1620
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
3216 AGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCGAGACTCAAACACTTGCTG 3275

1621 GGACCATCCCAAAATGACAGAGCTCACCAGTCTTTAGCTGACCTGGAATAATGTCAGATT 1680
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3276 GGACCATCCCAAAATGACAGAGCTCACCAGTCTTTAGCTGACCTGGAATAATGTCAGATT 3335

1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAGAACTGCGAGAGGCCCCTTGCTTGGATCT 1740
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3336 CTCAGCTTATAGGACTGCCATGAAACTCCGAGAACTGCGAGAGGCCCCTTGCTTGGATCT 3395

1741 CTTGAGCCTGTCAGCTGCATGTGATGCTTGATGCCTTGGACCAGCACAACCTCAAGCAAATGACCA 1800
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3396 CTTGAGCCTGTCAGCTGCATGTGATGCTTGATGCCTTGGACCAGCACAACCTCAAGCAAATGACCA 3455

1801 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGCCGCCTGGAGCA 1860
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3456 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGCCGCCTGGAGCA 3515

1861 AGAGCAACAACAATTGGTCAACGTCCCTCTCTGCGTGGATATGTCTGAACTGGCTGCT 1920
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3516 AGAGCAACAACAATTGGTCAACGTCCCTCTCTGCGTGGATATGTCTGAACTGGCTGCT 3575

1921 GAATGTTTATGATACGGACGAACAGGGAGATCCGTGTCCTGTCTCTTTTAAAACTGGCAT 1980
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3576 GAATGTTTATGATACGGACGAACAGGGAGATCCGTGTCCTGTCTCTTTTAAAACTGGCAT 3635

1981 CATTTCCCTGTGTAAAGCACATTTGGAAGAACAGACAGTAGATACCTTTTCAAGCAAGTGGC 2040
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3636 CATTTCCCTGTGTAAAGCACATTTGGAAGAACAGACAGTAGATACCTTTTCAAGCAAGTGGC 3695
```

```
  y    2041 AAGTTCAACAGGATTTTGTGACCAGCCGAGGCTGGCCTTCCTTCGCATGATTCTATCCA 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    3696 AAGTTCAACAGGATTTTGTGACCAGCCGAGGCTGGCCTTCCTTCGCATGATTCTATCCA 3755

  y    2101 A 2101
            |
  b    3756 A 3756
```

ESULT 3
S-10-149-736-40

Sequence 40, Application US/10149736
Publication No. US20030216332A1
GENERAL INFORMATION:
APPLICANT:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: Patentin version 3.1
SEQ ID NO 40
LENGTH: 5339
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
S-10-149-736-40

```
Query Match        83.8%;   Score 1760.4;   DB 15;   Length 5339;
Best Local Similarity   90.3%;   Pred. No. 0;
Matches 1910;   Conservative   0;   Mismatches 191;   Indels 15;   Gaps  2;
```

```
  y       1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACGACGGAGCCCATT   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACGACGGAGCCCATT 1158

  y      61 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG   120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    1159 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1218

  y     121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC   180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1278

  y     181 TGCTGAGGACAACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    1279 TGCTGAGGACAACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1338

  y     241 CCAGTTCATATCATGAGGGGTACATGATGGAAGTGCATCAGGGGCCATCAGGGCCGGGTTGG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    1339 CCAGTTCATATCATGAGGGGTACATGATGGAAGTGCATCAGGGGCCATCAGGGCCGGGTTGG 1398

  y     301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

  y     361 AACTGAAGTACAAGAGCAGCAGATGAATCCTAAATTCAAGATGGGAAATGGGAATGCCTAGC  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    1459 AACTGAAGTACAAGAGCAGCAGATGAATCCTAAATTCAAGATGGGAAATGGGAATGCCTAGC 1518

  y     421 TAGCATGATGAAAAACAAAGCAATTTACATAGAGTTTTAAGATCTGCAGAATCAGAAACT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
  b    1519 TAGCATGATGAAAAACAAAGCAATTTACATAGAGTTTTAAGATCTGCAGAATCAGAAACT 1578

  y     481 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAGAACAAGGAAAATGGAGGAAGA  540
            ||||  ||||  |||||    ||||
  b    2659 GGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTATATCAACCACGAGAC 2718
```

```
  Db   1579 GGAAAAGTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAATGTCCTACAGGATGC 1638

  Qy    541 GCCTCTTGGACCTGATCTTGAGACCTAAAAGGCCAGTACAACAACATAAGGTGCTTCA  600
            |||  |||||||
  Db   1639 TACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATG 1698

  Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT  660
            ||||| ||||  |||   ||  ||||
  Db   1699 GCAAGAACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAA 1758

  Qy    661 TGATGA---------ATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACT  708
            |||||
  Db   1759 CAGCCAAAAAATCCTGAGATCCCCGGAAGGTTCCGATGATGCAGTCGTCCTGTTACAAAGACG 1818

  Qy    709 TAAGG---TATTGGGAGATCGATGGCAAACATCTGTAGATGGACAGAACGACCGCTGGGT  765
            ||| |  ||||||||
  Db   1819 TTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTC 1878

  Qy    766 TCTTTTACAGACAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCT  825
            ||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   1879 CCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCT 1938

  Qy    826 GGTGTGGCTACAGCTGAAAGATGAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGACTT  885
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   1939 GGTGTGGCTACAGCTGAAAGATGAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGACTT 1998

  Qy    886 TCCAGCAGTTCAGAAGCAGCAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAA  945
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   1999 TCCAGCAGTTCAGAAGCAGCAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAA 2058

  Qy    946 AGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT 1005
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2059 AGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT 2118

  Qy   1006 GGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTGCTCCTCCTGAGGAGAGAGCCCA 1065
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2119 GGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTGCTCCTCCTGAGGAGAGAGCCCA 2178

  Qy   1066 GAATGTCACTCGGCTTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATT 1125
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2179 GAATGTCACTCGGCTTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAATT 2238

  Qy   1126 GAACCTGCACTCCCGCTGACTGGCAGAGAAAAATAGATGGGAGAAACTTGAAAGACTTCCAGGA 1185
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2239 GAACCTGCACTCCCGCTGACTGGCAGAGAAAAATAGATGAGGAGAAACTTGAAAGACTTCCAGGA 2298

  Qy   1186 ACTTCAAGAGAGGCCACGATGAGCTGGACCTCAAGCTGAGGTGATCAAGGG 1245
            |||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2299 ACTTCAAGAGAGGCCACGATGAGCTGGACCTCAAGCTGCGGTGATCAAGGG 2358

  Qy   1246 ATCCTGCAGCCCGTGGGCCGCCATCTCCATTGACTCTCCAAGATCACCTGAGAAAGT 1305
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2359 ATCCTGGCAGCCCGTGGGCCGCCATCTCCATTGACTCTCCAAGATCACCTGAGAAAGT 2418

  Qy   1306 CAAGGCACTTCGAGAGAATTGCCTCTGAAAAGAACGTGAGCCACGTCAATGACCT 1365
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2419 CAAGGCACTTCGAGAGAATTGCCTCTGAAAAGAACGTGAGCCACGTCAATGACCT 2478

  Qy   1366 TGCTCGGCCAGTTACCACTTTGGGCATTTGGGCATTTATCAGCTCTCAGCACTCTGA 1425
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2479 TGCTCGGCCAGTTACCACTTTGGGCATTTGGGCATTTATCAGCTCTCAGCACTCTGA 2538

  Qy   1426 AGACCTGAACCAGATGAAGCTTCTGCAGGTGGCCGTGCAGGACCAGGTCAGGCAGCT 1485
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2539 AGACCTGAACCAGATGAAGCTTCTGCAGGTGGCCGTGCAGGACCAGGTCAGGCAGCT 2598

  Qy   1486 GCATGAGCCCACAGGGACTTTGGTCAGATCTCCAGATCCAGATCTCTGTCCA 1545
            ||||||||||||||||||||||||||||||||||||||||||||||
  Db   2599 GCATGAGCCCACAGGGACTTTGGTCCAGATCTCAGCACTTTCTTTCCACGTCGTCCA 2658

  Qy   1546 GGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTATATCAACCACGAGAC 1605
            |||||||||||||||||||||||||||||||||||||||||||||||||
  Db   2659 GGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTATATCAACCACGAGAC 2718
```

```
RESULT 4
3-09-845-416-10
Sequence 10, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
   APPLICANT: XIAO, XIAO
   TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
   TITLE OF INVENTION: THEREOF
   FILE REFERENCE: DE1142
   CURRENT APPLICATION NUMBER: US/09/845,416
   CURRENT FILING DATE: 2001-04-30
   PRIOR APPLICATION NUMBER: 60/200,777
   PRIOR FILING DATE: 2000-04-28
   NUMBER OF SEQ ID NOS: 36
   SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
   LENGTH: 3531
   TYPE: DNA
   ORGANISM: Homo sapiens
3-09-845-416-10

Query Match      83.3%;    Score 1750.8;    DB 10;    Length 3531;
Best Local Similarity   90.2%;   Pred. No. 0;
Matches 1914;   Conservative   0;   Mismatches 187;   Indels 21;   Gaps 3;
```

(Page consists of paired DNA sequence alignment blocks labeled Qy (query) and Db (database), with position numbers 1 through 3531 and 181 through 3214.)

```
;   PRIOR APPLICATION NUMBER: 60/200,777
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 30
;     LENGTH: 4498
;     TYPE: DNA
;     ORGANISM: Homo sapiens
US-09-845-416-30

Query Match        83.3%;   Score 1750.8;   DB 10;   Length 4498;
Best Local Similarity   90.2%;   Pred. No. 0;
Matches 1914;   Conservative  0;   Mismatches 187;   Indels 21;   Gaps  3;

Qy      1 GAGTATGCCTACACACAGGCTGCTTATGTCACCACTCTGACCCTACACGGAGCCCATT    60
          ||| |||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1716

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG   120
          ||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||
Db   1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  1776

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGCTCTTTC   180
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGACCGTTATCAAACAGCTTTTAGAGAAGAAGTATTATCGTGGCTCTTTC 1836

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA  1896

Qy    241 CCAGTTCATACTCATGAGGGTACATGGGTACATGGAATTGACAGCCCATCAGGGGCCGGTTGG   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTCATACTCATGAGGGTACATGGGTACATGGAATTGACAGCCCATCAGGGGCGGGTTGG 1956

Qy    301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA 2016

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

Qy    421 TAGCATGGAAAAACAAAGCAATTACATAGAGTT------TTAATGGATCTCCAGAATCA   474
          ||||||||||||||||||||||||||||||||      ||||||||||||||||||||||
Db   2077 TAGCATGGAAAAACAAAGCAATTACATAGAACTCATAGATTACTGCAACAGTTCCCCCT  2136

Qy    475 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGACAAGAAAAATGGA   534
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2137 GGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAGACCTGAAGACTAAAACGCCAAGTGA  2196

Qy    535 GGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAGGT   594
          |||||       ||||
Db   2197 GGATGCTACCCGTAAGGAAGGGCTCCTAGAAGACTCCAGGGAGTAAAAAGAGCTGATGAA 2256

Qy    595 GCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGGT   654
          |||||||||||||
Db   2257 ACAATGGCAAGACTCCAAGGTGAAATTGAAGGTCACACAGATGTTTATCAAACCTGA   2316

Qy    655 GGTAG------------TTGATGAATCTAGTGGACATCAGCAACTGCTGTCTTTGGAAGA   702
                          ||||
Db   2317 TGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCGTTACA  2376

Qy    703 ACAACTTAAGG---TATTGGGAGATGGATCGGCAAACATCTGTAGGTGGACAAGAAGACCG   759
          ||||||   |||||||||||||
Db   2377 AAGAGGTTTGGATAACATGAACATTTCAAGTGGAGTGAACTTCGGAAAAAAGTCTCCAACAT 2436

Qy    760 CTGGGGTTCTTTACAGACAGTTCTGACCAGTGGAAGGCGTCTGCACCCTTTCTCTGCAGGA   819
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2437 CTGGGGTTCTTTACAGACAGTTCTGACCAGTGGAAGGCGTCTGCACCCTTTCTCTGCAGGA 2496

Qy    820 ACTTCTGGTGTGGCTACAGCTGAAAGATGAGAATGAATTAAGCCGGCAGCCATTGGAGG   879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2497 ACTTCTGGTGTGGCTACAGCTGAAAGATGAGAATGAATTAAGCCGGCAGCCATTGGAGG 2556
```

```
Qy 1240 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACCTCTCCAAGATCACCTCGA 1299
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2160 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACCTCTCCAAGATCACCTCGA 2219

Qy 1300 GAAAGTCAAGGCACTTCGAGGAGGAGAAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAA 1359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2220 GAAAGTCAAGGCACTTCGAGGAGGAGAAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAA 2279

Qy 1360 TGACCTTGCTCGGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 1419
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2280 TGACCTTGCTCGGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 2339

Qy 1420 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCCAGGACCGAGTCAG 1479
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2340 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCCAGGACCGAGTCAG 2399

Qy 1480 GCAGCTGCATGAGAGCCCACAGAGGACTTTGGTCCAGATCTCAGCACTTTCTTCCCACGTC 1539
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2400 GCAGCTGCATGAGAGCCCACAGAGGACTTTGGTCCAGATCTCAGCACTTTCTTCCCACGTC 2459

Qy 1540 TGTCCAGGGTCCCGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTATATCAACCA 1599
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2460 TGTCCAGGGTCCCGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTATATCAACCA 2519

Qy 1600 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 1659
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2520 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTTACCAGTCTTTAGC 2579

Qy 1660 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1719
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2580 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2639

Qy 1720 GAAGGGCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGAGTGCCTTGGACCAGCA 1779
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2640 GAAGGGCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGAGTGCCTTGGACCAGCA 2699

Qy 1780 CAACCTCAAGCAAAATGACCAGCCATGATCCTGCAGATATTCCTGAGATTATTGTTTGACCAC 1839
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2700 CAACCTCAAGCAAAATGACCAGCCATGATCCTGCAGATATTCCTGAGATTATTGTTTGACCAC 2759

Qy 1840 TATTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTTCGGTGGA 1899
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2760 TATTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTTCGGTGGA 2819

Qy 1900 TATGTGTCTGAAACTGGCTGCTGAAATGTTTATGATACGGGACGAACAGGGAGATCCGTGT 1959
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2820 TATGTGTCTGAAACTGGCTGCTGAAATGTTTATGATACGGGACGAACAGGGAGATCCGTGT 2879

Qy 1960 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGAAGTACAG 2019
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2880 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGAAGTACAG 2939

Qy 2020 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCCGCAGGCTGGGCCT 2079
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2940 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCCGCAGGCTGGGCCT 2999

Qy 2080 CCTTCTGCATGATTCATCCAA 2101
         |||||||||||||||||||||
Db 3000 CCTTCTGCATGATTCATCCAA 3021
```

```
Qy   1960 CCTGTCTTTAAAACTGGATCATCATTTCCCTGTGTAAAGCACATTTGAAGACAAGTACAG 2019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3637 CCTGTCTTTAAAACTGGATCATCATTTCCCTGTGTAAAGCACATTTGAAGACAAGTACAG 3696

Qy   2020 ATACCTTTCAAGCAAGTGCAAGTTCAACAGGATTTTGTGACCAGGATTTTGGGGCT 2079
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3697 ATACCTTTCAAGCAAGTGCAAGTTCAACAGGATTTTGTGACCAGCGCCAGGCTGGGGCT 3756

Qy   2080 CCTTCTGCATGATTCTATCCAA 2101
          ||||||||||||||||||||||
Db   3757 CCTTCTGCATGATTCTATCCAA 3778
```

RESULT 6
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9

Query Match   83.0%; Score 1743; DB 10; Length 3858;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

```
Qy      1 GAGCTATGCCTACACACAGCTGCTTATGTCACCACCTCTGACCACCACCTACACGGAGCCCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 GAGCCATGCCTACACACAGCTGCTTATGTCACCACCTCTGACCACCACGGAGCCCATT 959

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    960 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1019

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy    241 CCAGTTCATATCATGAGGGGTACATGATTGACAGCCATCAGGGCCGGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTCATATCATGAGGGGTACATGATTGACAGCCATCAGGGCCGGGTTGG 1199

Qy    301 TAATATTCTACAATTGGAAGTAAGCAGATAAGCTGATTGAACAGGAAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGAAGTAAGCAGATAAGCTGATTGAACAGGAAAAATTATCAGAAGATGAAGA 1259

Qy    361 AACTGAAGTACAAGAGCAAGATGAATCTCCTAAATTCAAGATGGGAATGGCTCAGGGTAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAAGATGAATCTCCTAAATTCAAGATGGGAATGGCTCAGGGTAGC 1319

Qy    421 TAGCATGGAAAACAAAGCAATTTACATGACGTTTAATGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 TAGCATGGAAAACAAAGCAATTTACATGACGTTTAATGAGTTTTAATGGATCTCCAGAATCAGAAACT 1379

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAACAACAAGGAAAAATGGGAGGAAGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAACAACAAGGAAAAATGGGAGGAAGA 1439
```

```
Qy   541  GCCTCTTGGACCTGATCTTGAAGAACGCCAAGTACAACAACATAAGGTGCTTCA         600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1440  GCCTCTTGGACCTGATCTTGAAGAACGCCAAGTACAACAACATAAGGTGCTTCA        1499

Qy   601  AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCTCACTCACGTGGTGGTAGT      660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1500  AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCTCACTCACGTGGTGGTAGT     1559

Qy   661  TGATGAATCTAGTGAGGATCACGCAACTGCTGTTGGAGAACAACTTAAGGTATTGGG     720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1560  TGATGAATCTAGTGAGGATCACGCAACTGCTGTTGGAGAACAACTTAAGGTATTGGG    1619

Qy   721  AGATGCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGA--  777
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1620  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC  1679

Qy   778  -------------------------------------------------------       777
Db  1680  TCATAGATTACTACTGCAACAGTTCCCCGTGGACCTGGAAAAGTTTCTGCCTGGCTTACAGA  1739

Qy   778  -------------------------------------------------------       777
Db  1740  AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA  1799

Qy   778  -------------------------------------------------------       777
Db  1800  CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAGACCTCCAAGGTGAAATTGAAGC  1859

Qy   778  -------------------------------------------------------       777
Db  1860  TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA  1919

Qy  1920  AGGTTCCGATGATGCAGTCCTGTTACAAAAGACGTTTGGAGTAACATGACTTCAAGTGGAG  1979
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   778  -----------------------------------------CAGTTCTGACCAGTG        792

Qy  1980  TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCCATTTGGAAGCCAGTTCTGACCAGTG  2039
          ||||||||||||||||||||||||||||||||||||||||
Db   778  -----------------------------------------------CAGTTCTGACCAGTG  792

Qy   793  GAAGCGTCGTCGCACCTTTCTTCTGCGCAGGAACTTCTGTGGTGTGCTACAGCTGAAAGATGATGA   852
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2040  GAAGCGTCGTCGCACCTTTCTTCTGCGCAGGAACTTCTCGGTGTGCTACAAGCTGAAAGATGATGA  2099

Qy   853  ATTAAGCGGCGGCAGGCGCACCTATTGGAGGCGACTTCAGCAGTTCAGAAGCAGAACGATGT   912
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2100  ATTAAGCGGCCGGCAGGCGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAGCAGAACGATGT  2159

Qy   913  ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA   972
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2160  ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA  2219

Qy   973  GACTGTACGAATATTTCTGACAGAGAGCTGGCGCTTCACTCGGCTTCGCTTCTACCAGAA  1032
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2220  GACTGTACGAATATTTCTGACAGAGAGCTGGCGCTTCACTCGGCTTCGCTTCTACCAGAA  2279

Qy  1033  GCCCAGAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCCTTCTACGAAAGCA  1092
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2280  GCCCAGAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCCTTCTACGAAAGCA  2339

Qy  1093  GGCTGAGGAGGTCAATACTGAGTGGGAAAAAAATTGAACCTGACACTGCCGCTACTGGCACAG  1152
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2340  GGCTGAGGAGGTCAATACTGAGTGGGAAAAAAATTGAACCTGACACTGCCGCTACTGGCACAG  2399

Qy  1153  AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGATGAGCTGGA  1212
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2400  AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGATGAGCTGGA  2459

Qy  1213  CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT  1272
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2460  CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT  2519
```

```
Qy  1273  CATTGACTCTCTCCAAGATCACCTCCGAGATCAAGTCAAGGCACTTGGAGAGAAATTGCGCC  1332
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2520  CATTGACTCTCTCCAAGATCACCTCCGAGATCAAGTCAAGGCACTTGGAGAGAAATTGCGCC  2579

Qy  1333  TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCGCAGCTTACCACTTTGGGCAT  1392
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2580  TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCGCAGCTTACCACTTTGGGCAT  2639

Qy  1393  TCAGCTCTCACCGTATAACCTCAGCACTCTGGAGACCTGAACACCAGATGGAGGCTTCT  1452
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2640  TCAGCTCTCACCGTATAACCTCAGCACTCTGGAGACCTGAACACCAGATGGAGGCTTCT  2699

Qy  1453  GCAGGTGGCCGTGCGAGGAGGCAGCAGTCAGGCAG||||GCCTGATGAAGAGCCCACAGGAGACTTGGTCC  1512
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2700  GCAGGTGGCCGTGCGAGGAGGCAGCAGTCAGGCAGCTGAGCCCACAGGACTTGGTCC  2759

Qy  1513  AGCATCTCAGCACTTTCTTTCCACGTGTCTCCAGGTGTCCCAGGGTCCCTGGGAGGAGAGCCATCTCCGCC  1572
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2760  AGCATCTCAGCACTTTCTTTCCACGTGTCTCCAGGTGTCCCAGGGTCCCTGGGAGGAGAGCCATCTCCGCC  2819

Qy  1573  AAACAAAGTGCCCTACTATATCAACCACCAGAACTCAAACAACTGCTCGGGACCATCCCAA  1632
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2820  AAACAAAGTGCCCTACTATATCAACCACCAGAGACTCAAACAACTGCTCGGGACCATCCCAA  2879

Qy  1633  AATGACAGAGCTCTACCAGTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG  1692
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2880  AATGACAGAGCTCTACCAGTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG  2939

Qy  1693  GACTGCCATGAAACTCCGAGAGACTGCAGAGGGCCCTTGCTTGGATCTCTTGAGCCTGTC  1752
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2940  GACTGCCATGAAACTCCGAGAGACTGCAGAGGGCCCTTTGCTTGGATCTCTTGAGCCTGTC  2999

Qy  1753  AGCTGCATGTGATGCCTTGCTGACCAGCACACCTCAAGCAAAAATGACCAGCCCATGGATAT  1812
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3000  AGCTGCATGTGATGCCTTGCTGACCAGCACACCTCAAGCAAAAATGACCAGCCCATGGATAT  3059

Qy  1813  CCTGCAGATTATTAATTGTTGACCTATTTATGACCAGAGACAACAA  1872
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  3060  CCTGCAGATTATTAATTGTTTGACCTATTTATGACCGCCTGAGCAAGAGACAACAACAA  3119

Qy  1873  TTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAACTGGCTGCTGAATGTTTATGA  1932
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3120  TTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAAACTGGCTGCTGAATGTTTATGA  3179

Qy  1933  TACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTAAAAACTGGCATCATTCCCTGTG  1992
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3180  TACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTAAAAACTGGCATCATTCCCTGTG  3239

Qy  1993  TAAAGCACATTGGAAGACAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG  2052
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3240  TAAAGCACATTGGAAGACAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG  3299

Qy  2053  ATTTTGTGACCAGCGCAGGCGTGGGCCTCCCTTCCTCTGCATGATTCTATCCAA  2101
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3300  ATTTTGTGACCAGCGCAGGCGTGGGCCTCCCTTCCTCTGCATGATTCTATCCAA  3348
```

```
RESULT 7
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
;   APPLICANT: XIAO, XIAO
;   TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;   TITLE OF INVENTION:    THEREOF
;   FILE REFERENCE: DE1142
;   CURRENT APPLICATION NUMBER: US/09/845,416
;   CURRENT FILING DATE:    2001-04-30
;   PRIOR APPLICATION NUMBER: 60/200,777
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
```

LENGTH: 4825
TYPE: DNA
ORGANISM: Homo sapiens
S-09-845-416-29

Query Match   83.0%;   Score 1743;   DB 10;   Length 4825;
Best Local Similarity   85.8%;   Pred. No. 0;
Matches 2101; Conservative   0; Mismatches   0; Indels 348; Gaps   1;

```
Qy        1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACCGAGCCCATT 60
            |||||||||| |||||||||||||||| |||  ||||||||||||||||||||||||||||
Db     1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACCGAGCCCATT 1716

Qy       61 TCCTTCACGCAGCATTGGAAGCTCCTGAGACAAGCAAGTCATTGGCAGTTCATTGATGAGAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1717 TCCTTCACAGCATTGGAAGCTCCTGAGACAAGCAAGTCATTGGCAGTTCATTGATGAGAG 1776

Qy      121 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     1777 TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

Qy      181 TGCTTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1837 TGCTTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy      241 CCAGTTCATATCTCATGAGGGTACATGATGATTGACAGCCCATCAGGGCGGTTGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1897 CCAGTTCATATCTCATGAGGGTACATGATGATTGACAGCCCATCAGGGCCGGTTGG 1956

Qy      301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1957 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

Qy      361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGCTTCAGGTAGC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2017 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGCTTCAGGTAGC 2076

Qy      421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2136

Qy      481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGA 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGA 2196

Qy      541 GCCTCTTGGAAGACCTGAAGTCTTGAAGACAAGACCTAAAAACGCCAAGTACAACACGGCCTTCA 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2197 GCCTCTTGGAAGACCTGAAGTCTTGAAGACAAGACCTAAAAACGCCAAGTACAACACATAAGGTGCTTCA 2256

Qy      601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCCACTCCACACTCGGAAGACTTAGT 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCCACTCCACATGTGTGGTAGT 2316

Qy      661 TGATGAATCTAGTGAGATCACGCAACATCTGAGATGGCAGAGAACCCGTCAGAGACTTGGG 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2317 TGATGAATCTAGTGAGATCACGCAACATCTGAGATGGCAGAGAACCCGTCAGAGACTTGGG 2376

Qy      721 AGATCGATGGGCAAACATCTGAAAACAAATGGTCTTCTTTACAAGA--- 777
            |||||||||||||||||||||||||||||||||||||||||||||
Db     2377 AGATCGATGGGCAAACATCTGAAAACAAATGGTCTTCTTTACAAGACAC 2436

Qy      778 ----------------------------------- 777
Db     2437 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCACCACGAGACTCAACAACAACTTGCTCGGGACCATCCCAA 2496

Qy      778 ----------------------------------- 777
Db     2497 AAACAAAGTGCCCTACTACTATCAACACCAAACAACTTGCTCGGGACCATCCCAA 2556

Qy      778 ----------------------------------- 777
Db     2557 CTCCAAGGGGAGTAAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2616

Qy      777 -------------------------- 777
Db     2617 TCACACAGATGTTTATCAACCTGGATGAAAACAGCCAAAAAATCTGAGATCCCTGGA 2676

Qy      777 -------------------------- 777
Db     2677 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAG 2736

Qy      778 -----CAGTTCTGACCAGTG 792
                 |||||||||||||||
Db     2737 TGAACTTCGGAAAAGTCTCTCAACATTGGAAGGCTACAGCTGAAAGATGATGA 2796

Qy      793 GAAGCGTCTGCACCTTCTGCGAGAACTCTGGTGTGGCTACAGCTGAAAGCAGAAGCAGCGATGT 852
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2797 GAAGCGTCTGCACCTTCTGCGAGGAACTCTGCGTGTGGCTACAGCTGAAAGCAGCAGCAGCGATGA 2856

Qy      853 ATTAAGCCCGCAGCCACCTATTGGGAGGCGACTTCCAGCAGTTCAGAAGCAGAACGATGT 912
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2857 ATTAAGCCCGCAGCCACCTATTGGGAGGCGACTTCCAGCAGTTCAGAAGCAGAACGATGT 2916

Qy      913 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 972
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2917 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2976

Qy      973 GACTGTACGAAATATTTCTGACAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGA 1032
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2977 GACTGTACGAAATATTTCTGACAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGA 3036

Qy     1033 GCCCAGAGAGCTGCCTCCTGGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1092
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3037 GCCCAGAGAGCTGCCTCCTGGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 3096

Qy     1093 GGCTGAGGAGGTCAATACTGAGTGGAAAAATTGAACCTGACTCGACTGGCGAGAG 1152
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3097 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGACTCGACTGGCGAGAG 3156

Qy     1153 AAAAATAGATGAGAACCCTGAAAACTCAGGAAACTTCAAGAGAGTCAGGATCAGAAACT 1212
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3157 AAAAATAGATGAGAACCCTGAAAACTCAGGAAACTTCAAGAGAGTCAGGATCAGAAACT 3216

Qy     1213 CCTCAAGCTGCGCCAAGCTGAGTGATGATCAAGGAGTCTCGGCAGCCCGTGGGGCGATCTCCT 1272
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3217 CCTCAAGCTGCGCCAAGCTGAGTGATGATCAAGGAGTCTCGGCAGCCCGTGGGGCCATCTCCT 3276

Qy     1273 CATTGACTCTCTCCAAGATCACCTCGAGATCAAGTCACCTGACCTTGCTCGCGCAGCTTACCACTTTGGGCCC 1332
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3277 CATTGACTCTCTCCAAGATCACCTCGAGATCAAGTCACCTGACCTTGCTCGCGCAGCTTACCACTTTGGGCC 3336

Qy     1333 TCTGAAAGAAACGTGAGCCGTCAATGACCTTGCTCTGCGCCAGCTTACCACCAGATGGAAGCTTCAT 1392
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3337 TCTGAAAGAAACGTGAGCCGTCAATGACCTTGCTCTGCGCCAGCTTACCACCAGATGGAAGCCAT 3396

Qy     1393 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACTGAACACCAGATGAAGTGGAAGCTTCT 1452
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3397 TCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACTGAACACCAGATGAAGTGGAAGCTTCT 3456

Qy     1453 GCAGGTGGCCGTCGAGGAGTCAGGCAGTCAGGAGGCAGCCGTCGCCGTCGCATCTGGTCC 1512
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3457 GCAGGTGGCCGTCGAGGAGTCAGGCAGTCAGGCAGTCGCATGGGACTTGGTCC 3516

Qy     1513 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCACCACGAGACTCAACAACAACTTGCTCGCC 1572
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3517 AGCATCTCAGCACTTTCTTTCCACGTCTGTCCACCACGAGAGCCCATCTGCC 3576

Qy     1573 AAACAAAGTGCCCTACTACTATCAACACCCCAGAGTTTCTTGCCTGGCTTACAGA 1632
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3577 AAACAAAGTGCCCTACTACTATCAACACCCCAA 3636

Qy     1633 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGA 1692
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3637 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTGCTTGAATTCTCAGCTTATAG 3696

Qy     1693 CTCCAAGGGGAGTAAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1752
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1752 GACTGCCATGAAACTCCGAAGACTCGAGAAGGCCCCTTTGCTTGGATCTCTTGAGCCTGTC
```

```
Qy    301  TAATATTCTACAAATGGGAAGTAAGCTGATTGAACAGCAGGAAAATTATCGAAGATGAAGA  360
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1980  TAATATTCTACAAATGGGAAGTAAGCTGATTGAACAGCAGGAAAAATTATCGAAGATGAAGA  2039

Qy    361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
                ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db   2040  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  2099

Qy    421  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2100  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  2159

Qy    481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAGGAAAAATGGAGGAAGA  540
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2160  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAGGAAAAATGGAGGAAGA  2219

Qy    541  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  600
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2220  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  2279

Qy    601  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT  660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2280  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT  2339

Qy    661  TGATGAATCTAGTGGACATCGCAACTGCTGCTTTGGAAGACCGCTGGGTTCTTTACAAGA--  720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2340  TGATGAATCTAGTGGACATCGCAACTGCTGCTTTGGAAGACCGCTGGGTTCTTTACAAGACAC  2399

Qy    721  AGATCGATGGGCAAACATCTGAGATGACAGAGAACCGCTGGGTTCTTTACAGA--------  777
                ||||||||||||||||||||||||||||||||||||||||||||
Db   2400  AGATCGATGGGCAAACATCTGAGATGACAGAGAACCGCTGGGTTCTTTACAAGACAC  2459

Qy    778  --------------------------------------------  777
Db   2460  TCATATGATTACTGCACAGTTCCCCCTGGACCTGGAAAAGTTTCTGCCTGGCTTACAGA  2519

Qy    778  --------------------------------------------  777
Db   2520  AGCTGAAACAACTGCCAATGTCCTACCAGGATGCTACCCGTAAGGAAAGGCTCCCAGAAGA  2579

Qy    778  --------------------------------------------  777
Db   2580  CTTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGAACCTCCAAGGTGAAATTGAAGC  2639

Qy    778  --------------------------------------------  777
Db   2640  TCACACAGATGTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA  2699

Qy    778  --------------------------------------------  777
Db   2700  AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACAGAACTTCAAGTGGAG  2759

Qy    792  ----------------------------CAGTTCTGACCAGTG  792
                                               ||||||||||||
Db   2760  TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTG  2819

Qy    852  GAAGCGGTCGCACCTTTCTCTGCGAGGAACTTCGGTGTGGCTACAGCTGAAAGATGATGA  852
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2820  GAAGCGGTCGCACCTTTCTCTGCGAGGAACTTCGGTGTGGCTACAGCTGAAAGATGATGA  2879

Qy    912  ATTAAGCCGGCAGGCACCTATTGGAGCGCACTTCAGCAGTTCAGAAGCAGAACGATGT  912
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2880  ATTAAGCCGGCAGGCACCTATTGGAGCGCACTTCAGCAGTTCAGAAGCAGAACGATGT  2939

Qy    972  ACATAGGGCCTTCAAGAGGGATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA  972
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2940  ACATAGGGCCTTCAAGAGGGATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA  2999

Qy    1032 GACTGTACGAATATTTCTGACAGAGGACCTTGAAGAACTAGAGAAACTCTACCAGGA  1032
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3000  GACTGTACGAATATTTCTGACAGAGGACCTTGAAGAACTAGAGAAACTCTACCAGGA  3059

Qy    1092 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGCGCTTCACGAAAGCA  1092
```

```
Db   3697  GACTGCCATGAAACTCCGAAGACTGGACGCACAACCTCAAGCAAAATGACCAGCCATGGATAT  3756
Qy    1753 AGCTGCAGTGATGCCTTGGACCAGCAGATGTTTGACCACTATTGTTGACCAGAATCAGCACAA  1812
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3757  AGCTGCAGTGATGCCTTGGACCAGCAGATGTTTGACCACTATTGTTGACCAGAGCAGCACAA  3816

Qy    1813 CCTGCAGATTATTAATTGTTGACCAGTCGTCTGTCTCAAGCAAGAGCAAGAGCACAACAA  1872
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3817  CCTGCAGATTATTAATTGTTGACCAGTCGTCTGTCTCAAGCAAGAGCAAGAGCACAACAA  3876

Qy    1873 TTTGGTCAACGTCCCTCTCTGCGTGATATGTGTCTGAACTGGCTGCTGAAGTGTTATGA  1932
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3877  TTTGGTCAACGTCCCTCTCTGCGTGATATGTGTCTGAACTGGCTGCTGAAGTGTTTATGA  3936

Qy    1933 TACGGGACGAACAGCAGGAGGATCCGTGTCGTCCTGTCTTTAAAACTGGCATCATTCCCTGTG  1992
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3937  TACGGGACGAACAGCAGGAGGATCCGTGTCGTCCTGTCTTTAAAACTGGCATCATTCCCTGTG  3996

Qy    1993 TAAAGCACATTGGAGAACAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG  2052
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3997  TAAAGCACATTGGAGAACAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG  4056

Qy    2053 ATTTTGTGACCAGCGCCAGGCTGGGCCTCCTTCCTCTGCATGATTCTATCCAA  2101
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4057  ATTTTGTGACCAGCGCCAGGCTGGGCCTCCTTCCTCTGCATGATTCTATCCAA  4105
```

RESULT B

US-09-845-416-35
Sequence 35, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
  APPLICANT: XIAO, XIAO
  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
  TITLE OF INVENTION: THEREOF
  FILE REFERENCE: DE1142
  CURRENT APPLICATION NUMBER: US/09/845,416
  CURRENT FILING DATE: 2001-04-30
  PRIOR APPLICATION NUMBER: 60/200,777
  PRIOR FILING DATE: 2000-04-28
  NUMBER OF SEQ ID NOS: 36
  SOFTWARE: PatentIn Ver. 2.1
  SEQ ID NO 35
    LENGTH: 4848
    TYPE: DNA
    ORGANISM: Homo sapiens
US-09-845-416-35

Query Match   83.0%;   Score 1743;   DB 10;   Length 4848;
Best Local Similarity 85.8%;   Pred. No. 0;
Matches 2101;   Conservative 0;   Mismatches 0;   Indels 348;   Gaps 1;

```
Qy      1  GAGCTATGCCTACACAGGCTGCTTATGTCCATGTCACCACCTCTGACCCTACACGGAGCCCATT  60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1680  GAGCTATGCCTACACAGGCTGCTTATGTCCATGTCACCACCTCTGACCCTACACGGAGCCCATT  1739

Qy     61  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1740  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  1799

Qy    121  TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1800  TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  1859

Qy    181  TGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA  240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1860  TGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA  1919

Qy    241  CCAGTTTCATACTTCATGAGGGGTACATGAGGGGTACAGCCCATCAGGGCCGGGTTGG  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1920  CCAGTTTCATACTTCATGAGGGGTACATGAGGGGTACAGCCCATCAGGGCCGGGTTGG  1979
```

US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match   83.0%; Score 1743; DB 10; Length 5060;
Best Local Similarity   85.8%; Pred. No. 0;
Matches 2101; Conservative   0; Mismatches   0; Indels 348; Gaps   1;

Qy   1 GAGCTATGCCTACACACAGCGTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1892 GAGCTATGCCTACACACAGCGTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT 1951

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1952 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 2011

Qy   121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGACAGATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2012 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGACAGATTATCGTGGCTTCTTTC 2071

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2072 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 2131

Qy   241 CCAGTTCATACTCATGAGGGGTACATGATGAAGACAGCCCATCAGGGCCGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2132 CCAGTTCATACTCATGAGGGGTACATGATGAAGACAGCCCATCAGGGCCGGGTTGG 2191

Qy   301 TAATATTCTACAATTGGAAGTAAGCTGATTGATTGGAACACAGGAAAATTATCAGAAGAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2192 TAATATTCTACAATTGGAAGTAAGCTGATTGATTGGAACACAGGAAAATTATCAGAAGAA 2251

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2252 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 2311

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 2371

Qy   481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAACAGAAGAAATGAGGAGAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAACAGAAGAAATGAGGAGAGA 2431

Qy   541 GCCTCTTGACCTGATCTTGAAGACCCAAGTACAACAACATAAGGTGCTTCA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2432 GCCTCTTGACCTGATCTTGAAGACCCAAGTACAACAACATAAGGTGCTTCA 2491

Qy   601 AGAAGACTAGAACAAGAACAAGTCAGGGTCAATTCTCCACATGGTCGTGTGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2492 AGAAGACTAGAACAAGAACAAGTCAGGGTCAATTCTCCACATGGTCGTGTGTAGT 2551

Qy   661 TGATGAATCTAGTGAGTGGAGATCACGCAACTGCTGTTGGAAGAACAACTAAGTATTGGG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2552 TGATGAATCTAGTGAGTGGAGATCACGCAACTGCTGTTGGAAGAACAACTAAGTATTGGG 2611

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGTTCTTTTACAAGA-- 777

.b 3060 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         GCCCAGAGAGCTGCCTCCCTGAGGAGAGCCCAGAATGTGTTCACTCGGCTTCTACGAAAGCA 3119

.y 1093 GGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGACCTCCGCTGACTGGCAGAG 1152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3120 GGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGACCTCCGCTGACTGGCAGAG 3179

.y 1153 AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGACTCCACGGATGAGCTGGA 1212
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3180 AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGACTCCACGGATGAGCTGGA 3239

.y 1213 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGTGGGGCGATCCT 1272
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3240 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGTGGGGCCATCCT 3299

.y 1273 CATTGACTCTTCTCCAAGATCACTTCGAGAAAGTCAAGGCACTTCGAGGGAAATTGCGCC 1332
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3300 CATTGACTCTTCTCCAAGATCACTTCGAGAAAGTCAAGGCACTTCGAGGGAAATTGCGCC 3359

.y 1333 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCGCAGCTTACCACTTTGGGCAT 1392
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3360 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCGCAGCTTACCACTTTGGGCAT 3419

.y 1393 TCAGCTCTCACCGTATAACCTCAGCACTCGGAGAACTGAACCCAGATGGAAGTTTCT 1452
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3420 TCAGCTCTCACCGTATAACCTCAGCACTCGGAGAACTGAACCCAGATGGAAGTTTCT 3479

.y 1453 GCAGGTGGCCGTGCGAGGACCGAGTCAGGCGCAGTCAGGCGACTGAGAAGCCCACAGGGACTTTGGTCC 1512
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3480 GCAGGTGGCGTCGGAGGACCGAGTCAGGCGCAGTCAGGCGACTGAGAAGCCCACAGGGACTTTGGTCC 3539

.y 1513 AGCATCTCAGCACTTTCTTTTCACGTCTGTCCAGGGTCCCTGGGAGGAGCCATCTCGCC 1572
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3540 AGCATCTCAGCACTTTCTTTTCACGTCTGTCCAGGGTCCCTGGGAGGAGCCATCTCGCC 3599

.y 1573 AAAACAAAGTGCCCTACTATATCAACAACAGACTCAAACTTGCTGGACCATCCCAA 1632
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3600 AAAACAAAGTGCCCTACTATATCAACAACAGACTCAAACTTGCTGGACCATCCCAA 3659

.y 1633 AATGACAGAGCTCTACCAGTCTTAGCTGAGTCTTGACCTGAATAATGTCAGATTCTCAGCTTATAG 1692
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3660 AATGACAGAGCTCTACCAGTCTTAGCTGAGTCTTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3719

.y 1693 GACTGCCATGAAACTCTGCAGAAGACTTGCAGAAGGGCCCTTTGCTTGGATCTCTTGGAGCCTGTC 1752
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3720 GACTGCCATGAAACTCTGCAGAAGACTTGCAGAAGGGCCCTTTGCTTGGATCTCTTGGAGCCTGTC 3779

.y 1753 AGCTGCATGTGATGCGTCTTGGACGCAACAGCAAAATGACCAGCCCATGGATAT 1812
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3780 AGCTGCATGTGATGCCTTGGACGCAACAGCAAAATGACCAGCCCATGGATAT 3839

.y 1813 CCTGCAGATTATTAATTGTTGACCACTATTATGACCGCCTGAGCAAGAGCACAACAA 1872
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3840 CCTGCAGATTATTAATTGTTGACCACTATTATGACCGCCTGAGCAAGAGCACAACAA 3899

.y 1873 TTTGGTCAACGTCCCTCTCTGCGCGTGGATATGTGTCTGAATGTGCTCTGAAATGTTTATGA 1932
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3900 TTTGGTCAACGTCCCTCTCTGCGCGTGGATATGTGTCTGAATGTGCTCTGAAATGTTTATGA 3959

.y 1933 TACGGGACGAACAGGGAGGAGGATCCGTCCGTGTCCTCGTCTTCCTGTCTCTTTAAAACTGGCATCATTTCCCTGTG 1992
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 3960 TACGGGACGAACAGGGAGGAGGATCCGTCCGTGTCCTCGTCTTCCTGTCTCTTTAAAACTGGCATCATTTCCCTGTG 4019

.y 1993 TAAAGCACATTGAAGACAAGCAAGTACAGATACCTTTTCAAGCAAGTGCCAAGTTCAACAGG 2052
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 4020 TAAAGCACATTGGAAGACAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4079

.y 2053 ATTTTGTGACGCCAGCCAGCGCAGGGCCTGGGCCTCCTTCCTTCTGCATGAATTCTATCCAA 2101
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b 4080 ATTTTGTGACCAGCCAGCGCAGGGCCTGGGCCTCCTTCCTTCTGCATGAATTCTATCCCAA 4128

RESULT 9

RESULT 10
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US2003017131A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match   78.1%; Score 1641; DB 10; Length 4414;
Best Local Similarity   88.0%; Pred. No. 0;
Matches 1848; Conservative   0; Mismatches 190; Indels 63; Gaps 3;

```
Qy      1 GAGCTATGCTACACAGGCTGCTTATGTCACCACTTCGACCCTACACGGAGCCCATT   60
          ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db   1657 GAGCTATGCTACACAGGCTGCTTATGTCACCACTTCGACCCTACACGGAGCCCATT 1716

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAG  120
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1776

y     121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

y     181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

y     241 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    1897 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1956

y     301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

y     361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

y     421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2136

y     481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA 2196

y     541 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256

y     601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 2316

y     661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376

y     721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 2436

y     781 TTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCT 840
             |  | | ||| | |||||          | |       ||   |  |||||
b    2437 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 2496

y     841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAA 900
           ||||| |||| | | | |  |       | | |||     |   || |   |    |
b    2497 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 2556

y     901 GCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCAT 960
              |    |  | |   |||| |    |  |       |   ||  ||     |  |||
b    2557 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2616

y     961 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 1020
          | | |  ||  |              |||   || |  ||   ||      ||     |
b    2617 GGGCAAACTGTA--------------TTCACTCAAACAAGATCTTCTTTCAACACTGAAG 2662

y    1021 ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCT 1080
          | |    |||   ||||||   | |   | ||    || | |   |   |    | ||
b    2663 AATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGG-- 2720

y    1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGC 1140
                                   |  | | |||   |||| |||||
b    2721 -------------------------ATAATTTAGTCCAAAAACTTGAA------------ 2743

y    1141 TGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 1200
                    |  |  | ||   |||||||||||||||||||||||||||||||||||||
Db   2744 ----------AAGAGTACAGCACAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCAC 2793

Qy   1201 GGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2794 GGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 2853

Qy   1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2854 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 2913

Qy   1321 AGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2914 AGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 2973

Qy   1381 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2974 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 3033

Qy   1441 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3034 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAG 3093

Qy   1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3094 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 3153

Qy   1561 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3154 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 3213

Qy   1621 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3214 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 3273

Qy   1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3274 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCT 3333

Qy   1741 CTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3334 CTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCA 3393

Qy   1801 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3394 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA 3453

Qy   1861 AGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3454 AGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCT 3513

Qy   1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3514 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCAT 3573

Qy   1981 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3574 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 3633

Qy   2041 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTCTATCCA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3634 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTCTATCCA 3693

Qy   2101 A 2101
          |
Db   3694 A 3694

RESULT 11
US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
;  APPLICANT: XIAO, XIAO
```

TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
  LENGTH: 3446
  TYPE: DNA
  ORGANISM: Homo sapiens
IS-09-845-416-14

Query Match   77.5%;   Score 1629; DB 10;   Length 3446;
Best Local Similarity   87.9%;   Pred. No. 0;
Matches 1847;   Conservative 0;   Mismatches 190;   Indels 64;   Gaps 4;

```
Db  1679  CCTTCCAAATGGCAACGTCTTACTGACGAAGACAGTGCCTTTTAGTGCATGGCTTTCAGA  1738
Qy   841  GAAAGATGATGAATTAAGCCGGCAGCGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAA   900
Db  1739  AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT  1798
Qy   901  GCAGAACGATGTACATAGGGCCTTCAAGAGGGATTGAAAACTAAAGAACCTGTAATCAT   960
Db  1799  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAAGAAAAAAAGCAATCCAT  1858
Qy   961  GAGTACTTCTGTGAGACTGTACGAGAATATTCTGACAGAGAGCCTTTGGAAGGACTAGAGAA  1020
Db  1859  GGGCAAACTGTA---------TTCACTCAAACAACAGATCTTCTTTCAACACTGAAG  1904
Qy  1021  ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCT  1080
Db  1905  AATAAGTCAGTGACCCAGAAGACGAAGCATGGCTGGA---------  1942
Qy  1081  TCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGC  1140
Db  1943  ----TAACTTTGCCCGGTGTGTGGGATAATTTAGTCAAAACTTGAA---------  1985
Qy  1141  TGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTTCAAGAGACTTCAAGAGGGCCAC  1200
Db  1986  ---------AAGAGTACAGCACAGACCCTCAGCGTGCGCGCAAGCTGCGCGCAAGCTGCGCCCAC  2035
Qy  1201  GGATGAGCTGACCTCAAGCTGCGCGCAAGCTGCGCGCAAGCTGAGTGATCAAGGATCCTGGCAGCCCGT  1260
Db  2036  GGATGAGCTGGACCTCAAGCTGCGCGCAAGCTGCGCGCAAGCTGAGTGATCAAGGATCCTGGCAGCCCGT  2095
Qy  1261  GGGCGATCTCTCTCATTGACTTCTCCAAGATCACCTCAAGGAACTCGAGG  1320
Db  2096  GGGCGATCTCCCTCATTGACTTCTCCAAGATCACCTCGAGAAAGTCAAGGACTTCGAGG  2155
Qy  1321  AGAAAATTGCGCCTCTGAAAAGAACGTGAGCCACGTCAATGACCTGCTCGCCAGCTTAC  1380
Db  2156  AGAAAATTGCGCCTCTGAAAAGAACGTGAGCCACGTCAATGACCTGCTCGCCAGCTTAC  2215
Qy  1381  CACTTTGGGCATTCAGCTCCACGTATAACCTGCACCTCTCTGAAGACCTGAACACCAG  1440
Db  2216  CACTTTGGGCATTCAGCTCCACGTATAACCTGCACCTCTCTGAAGACCTGAACACCAG  2275
Qy  1441  ATGGAGCTTCTGCAGGTGGCCGTCGAGGACCGAGCTGCATGAAGCCCACAG  1500
Db  2276  ATGGAGCTTCTGCAGGTGGCCGTCGAGGACCGAGCTGCATGAAGCCCACAG  2335
Qy  1501  GGACTTGGTCGCAGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCTGGGAGAG  1560
Db  2336  GGACTTGGTCGCAGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCTGGGAGAG  2395
Qy  1561  AGCCATCTCGCCAACAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG  1620
Db  2396  AGCCATCTCGCCAACAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG  2455
Qy  1621  GGACCATCCCAAAATGACAGAGCTCTTAGCTGACCTGAATAATGTCAGATT  1680
Db  2456  GGACCATCCCAAAATGACAGAGCTCTTAGCTGACCTGAATAATGTCAGATT  2515
Qy  1681  CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCGAGAGGCCCTTTGCTTGGATCT  1740
Db  2516  CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCGAGAGGCCCTTTGCTTGGATCT  2575
Qy  1741  CTTGAGCCTGTCAGCTGCATGTGATGCCTTGGAACCAGCAGCACAACCTCAAGCAAATGACCA  1800
Db  2576  CTTGAGCCTGTCAGCTGCATGTGATGCCTTGGAACCAGCAGCACAACCTCAAGCAAATGACCA  2635
Qy  1801  GCCCATGATCTGAGATTATTAATTGTTGACCACTATTATGACCGCCTGGAGACAG  1860
Db  2636  GCCCATGATCTGAGATTATTAATTGTTGACCACTATTATGACCGCCTGGAGACA  2695
Qy  1861  AGAGCACAACAATTGGTCAACGTCCCTCTGCGTGTGGGTATGTGTCTGCTGCT  1920
Db  2696  AGAGCACAACAATTGGTCAACGTCCCTCTGCGTGTGGGTATGTGTCTGCTGCT  2755
```

```
Qy     1  GAGCTATGCCTACACAGAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT    60
Db   900  GAGCTATGCCTACACAGAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   959
Qy    61  TCCTTCACAGCATTTGGAAGCTCCTGAAGACACAGTCGAGTTCATTGGCAGTTCATTGATGGAGAG   120
Db   960  TCCTTCACAGCATTTGGAAGCTCCTGAAGACACAGTCGAGTTCATTGGCAGTTCATTGATGGAGAG  1019
Qy   121  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
Db  1020  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  1079
Qy   181  TGCTGAGGACACATTGCAAGCACAACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA   240
Db  1080  TGCTGAGGACACATTGCAAGCACAACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA  1139
Qy   241  CCAGTTTCATCATGAGGGTACATGATTGGATTGACAGCCCATTGACAGGGCCCGGGTTGT   300
Db  1140  CCAGTTTCATCATCAGAGGGTACATGATTGGATTGACAGCCCATTGACAGGGCCCGGGTTGG  1199
Qy   301  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   360
Db  1200  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  1259
Qy   361  AACTGAAGTACAAGAGAGCAGAGATGAATCTCCTAAAATTCAAGATGGGAAATGCTCAGGGTAGC   420
Db  1260  AACTGAAGTACAAGAGAGCAGAGATGAATCTCCTAAAATTCAAGATGGGAAATGCTCAGGGTAGC  1319
Qy   421  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAAACT   480
Db  1320  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATC-GAAACT  1378
Qy   481  GAAAGAGTTGAATGACTGGCTAACAAAACAAGGAAGCAAGGAAATGGAGGAAGA   540
Db  1379  GAAAGAGTTGAATGACTGGCTAACAAAACAAGGAAGCAAGGAAATGGAGGAAGA  1438
Qy   541  GCCTCTTGGACCTGATCTTGAAGACAATCGACTCGAGAAAGTGGAAGACCTTAAGGTGCTTCA   600
Db  1439  GCCTCTTGGACCTGATCTTGAAGACAATCGACTCGAGAAAGTGGAAGACCTTAAGGTGCTTCA  1498
Qy   601  AGAAGATCTAGAACAAGACAAGAGACTCAGGGTCAATTCTCTACCTCACATGGTGGTAGT   660
Db  1499  AGAAGATCTAGAACAAGACAAGAGACTCAGGGTCAATTCTCTCACTCACATGGTGGTAGT  1558
Qy   661  TGATGAATCTAGTGGACTCAGCTCAGCCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG   720
Db  1559  TGATGAATCTAGTGGACTCAGCTCAGCCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG  1618
Qy   721  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACACCGGCTGGGTTCTTTTACAAGACAG   780
Db  1619  AGATCGATGGGCAAACATCTGTGTAGATGGACAGAAGACACCGCTGGGTTCTTTTACAGACAT  1678
Qy   781  TTCTGACCAGTGGAAGCGTCTGCACCTTCTGCTGCGCCAGGAACTTCGTGTGTGGCTACAGCT   840
Db
```

Db 1459 AACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGATGTGGAATGCCTCAGGGTAGC 1518
Qy 421 TAGCATGAAAAACAAAGCAATTACATAGAGTT----------------------------------- 454
Db 1519 TAGCATGAAAAACAAAGCAATTTACATGTCCTCGACCTGACCACTATTGGAGCCTCCC 1578

Qy 455 --- 454
Db 1579 TACTCAGACTGTTACTCTGGTGCACACAACCTTTGGTGGTTACTAAGGAAACTGCCATCTCCAA 1638

Qy 455 ---TTAATGGATCTCCAGAATCAGAA 477
Db 1639 ACTAGAAATGCCATCTTCCTTGATGTTGGAGCATAGATTACTGCAACAGTTCCCCTGGA 1698

Qy 478 ACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAGGAAAATGGAGGA 537
Db 1699 CCTGGAAAAGTTTCTTGCCTGGCTTACAGAGACCTGAAACAACTGCCAATGTCCTACAGGA 1758

Qy 538 AGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACACAACATAAGGTGCT 597
Db 1759 TGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACA 1818

Qy 598 TCAAGAGATCTAGAACAAGAACAAGTCAGGGTGCAATTCTCCCACTCCACATGGTGGTGGT 657
Db 1819 ATGGCAAGACCCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGA 1878

Qy 658 AGTTGATGA-----------ATCTAGTGGAGATCACGCAACTGCTGTTTGGAAGAACA 705
Db 1879 AAACAGCCAAAAAAATCCTGAGATCCCTGAAGGTTCCGATGATGCAGTCCTGTTACAAAG 1938

Qy 706 ACTTAAGG---TATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGCTG 762
Db 1939 ACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAG 1998

Qy 763 GGTTCTTTTACAGACAGTTCTGACCAGTGAAGCGTCTGCACCTTTCTCTGCAGGGAACT 822
Db 1999 GTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGGAACT 2058

Qy 823 TCTGGTGTGGCTACAGCTGAAAAGTGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGA 882
Db 2059 TCTGGTGTGGCTACAGCTGAAAAGTGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGA 2118

Qy 883 CTTTCCAGCAGTTCAGAAGCAGCAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAC 942
Db 2119 CTTTCCAGCAGTTCAGAAGCAGCAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAC 2178

Qy 943 TAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCGACAGCAGCGCC 1002
Db 2179 TAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCGACAGCAGCGCC 2238

Qy 1003 TTTGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTCGCCTCCTGAGGAGAGAGC 1062
Db 2239 TTTGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGC 2298

Qy 1063 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAA 1122
Db 2299 CCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAA 2358

Qy 1123 ATTGAACCTGCCACTCCGCTGACTCGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCA 1182
Db 2359 ATTGAACCTGCCACTCCGCTGACTCGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCA 2418

Qy 1183 GGAACTTCAAGAGCCACGAGGCCTGACCTGAGCTGGACCTGAAAGCTGCGCCAAGCTGAGGTGATCAA 1242
Db 2419 GGAACTTCAAGAGCCACGAGGCCTGACCTGAGCTGGACCTGAAAGCTGCGCCAAGCTGAGGTGATCAA 2478

Qy 1243 GGGATCCTGGCAGCCCGTGGGGCGATCTCGGCTTCCTCATTGACTTCCTCATTGACTGGCTTCTTTC 1302
Db 2479 GGGATCCTGGCAGCCCGTGGGGCGATCTCGGCTTCCTCATTGACTTCCTCATTGACTGGCTTCTTTC 2538

Qy 1303 AGTCAAGGCACTTCGAGGAATTGCGCCCTCTGAAAGAAGAACGTGAGCCACGTCAATGA 1362
Db 2539 AGTCAAGGCACTTCGAGGAATTGCGCCCTCTGAAAGAAGAACGTGAGCCACGTCAATGA 2598

Qy 1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTCCGTCGTCTTTAAAACTGGCAT 1980
Db 2756 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTCCGTCGTCTTTAAAACTGGCAT 2815

Qy 1981 CATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTCAAGCAAGTGGC 2040
Db 2816 CATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2875

Qy 2041 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTCTATCCA 2100
Db 2876 AAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGCATGATTCTATCCA 2935

Qy 2101 A 2101
Db 2936 A 2936

RESULT 12
US-10-149-736-41
Sequence 41, Application US/10149736
Publication No. US20030216332A1
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41
LENGTH: 5462
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-149-736-41

Query Match　　77.5%;　　Score 1629;　DB 15;　Length 5462;
Best Local Similarity　85.4%;　　Pred. No. 0;
Matches 1911;　Conservative　0;　Mismatches 190;　Indels 138;　Gaps 3;

Qy   1 GAGCTATGCCTACACACAGCGCTGCTTATGTCACCACCTCTGACCCTACACCGGAGCCCATT 60
Db 1099 GAGCTATGCCTACACACAGCGCTGCTTATGTCACCACCTCTGACCCTACACCGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
Db 1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

Qy 121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 1278

Qy 181 TGCTGAGGACACATTGCAAGCACAAGGAGCACAAGGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAGGAGCACAAGGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy 241 CCAGTTTCATACTCATCATGAGGGGTACATGATGGGTACATGATGGGTACAGCGCCCATCAGGGCCCGGGGTTGG 300
Db 1339 CCAGTTTCATACTCATCATGAGGGGTACATGATGGGTACATGATGGGTACAGCGCCCATCAGGGCCCGGGGTTGG 1398

Qy 301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db 1399 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458

Qy 361 AACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
Db   AACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC

```
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-10-149-736-39

Query Match             76.7%;  Score 1611.6;  DB 15;  Length 5417;
Best Local Similarity   85.3%;  Pred. No. 0;
Matches 1872;  Conservative  0;  Mismatches  229;  Indels  93;  Gaps  3;

Qy      1  GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCTACACGGAGCCCATT    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099  GAGCTATGCTACACACAGGCTGCTTATGTCAACCACCTCTGACCCTACACGGAGCCCATT  1158

Qy     61  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  1218

Qy    121  TGAAGTAAACCTGGACCGTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219  TGAAGTAAACCTGGACCGTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC  1278

Qy    181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  1338

Qy    241  CCAGTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339  CCAGTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG  1398

Qy    301  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAAGGAAAATTATCAGAAGATGAAGA   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAAGGAAAATTATCAGAAGATGAAGA  1458

Qy    361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGCCTCAGGGTAGC   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGCCTCAGGGTAGC  1518

Qy    421  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  1578

Qy    481  GAAAGAGTTGAATGACTGGCTAACAAAAACGAAGAAGAAGAACAAGAAAATGGAGGAAGA   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579  GAAAGAGTTGAATGACTGGCTAACAAAAACGAAGAAGAAGAACAAGAAAATGGAGGAAGA  1638

Qy    541  GCCTCTTGGACCTGATCTTGAGACCTCAAAAACGCCAAGTACAACAACATAAGGTGCTTCA   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639  GCCTCTTGGACCTGATCTTGAGACCTCAAAAACGCCAAGTACAACAACATAAGGTGCTTCA  1698

Qy    601  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCCTCACTCACATGGTGGTGGTAGT   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCCTCACTCACATGGTGGTGGTAGT  1758

Qy    661  TGTGAATCTAGTGGAATCTGGAGATCACGGACAACTGCTCGTCGTTTTACAGAGACAG   720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759  TGTGAATCTAGTGGAATCTGGAGATCACGGACAACTGCTCGTCGTTTTACAGAGACAG  1818

Qy    721  AGATCGATGGCAAACATCTGTAGATGACAGCAAGAACCGCTGGGTTCTTTTACAGAGACAT   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819  AGATCGATGGCAAACATCTGTAGATGACAGCAAGAACCGCTGGGTTCTTTTACAGAGACAT  1878

Qy    781  TTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGAGGAACTTCTCGGTGTGGCTACAGCT   840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879  CCTTCTCAAATGGCAACGTCCTTACTGAAGACAGTGCCTTTTTAGTGCATGGCTTTCAGA  1938

Qy    841  GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCGACTTTCCAGCAGTTCAGAA   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939  AAAAGAGATCAGTGAACAAGATTCACACACTGGCTTTAAAGATCAAAATGAAAATGTT  1998

Qy    901  GCAGAACGATGTACATAGGGCC--TTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATC   958
           ||||||||||||||||||||||||   |||
Db   1999  ATCAAGTCTTCAAAAACTGGCCGTGTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT  2058
```

RESULT 13
US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 5417

```
y   1363  CCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTTCACCGTATAACCTCAGCACTCT  1422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2599  CCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTTCACCGTATAACCTCAGCACTCT  2658

y   1423  GGAAGACCTGAACACACCAGATGGAAGCTTCTGCGAGGTGGCCGTCGAGGACCGAGTCAGGCA  1482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2659  GGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCCGTCGAGGACCGAGTCAGGCA  2718

y   1483  GCTGCATGAAGCCCACAGGGACTTTGGTCCAGCACTCTTCTTTCTTCACCGTCTGT  1542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2719  GCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCTTCACCGTCTGT  2778

y   1543  CCAGGGTCCCTGGGAGAGGAGCCATCTCGGCCCAAACAAAGTGCCCTACTATATCAACCACCGA  1602
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2779  CCAGGGTCCCTGGGAGAGGAGCCATCTCGGCCCAAACAAAGTCCCTACTATATCAACCACGA  2838

y   1603  GACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCCTCCACCAGTCTTAGCTGA  1662
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2839  GACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCCTCCACCAGTCTTTAGCTGA  2898

y   1663  CCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAATCCGAAGACTGCAGAA  1722
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2899  CCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAATCCGAAGACTGCGAGAA  2958

y   1723  GGCCCTTGCTTGGATCCTTGAGCCTGTCAGCTGCATGTGATGCCTGCTGGACCAGCACAA  1782
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2959  GGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTGCTGGACCAGCACAA  3018

y   1783  CCTCAAGCAAAATGACCAGCCCATGGGATAATCCTGCAGATTATTAATTGTTTGACCACTAT  1842
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3019  CCTCAAGCAAAATGACCAGCCCATGGGATAATCCTGCAGATTATTAATTGTTTGACCACTAT  3078

y   1843  TTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCCTCGCGTGGATAT  1902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3079  TTATGACCGGCTGGAGCAAGAGCAACAACAATTGGTCAACGTCCCTCCTCTCGCGTGGATAT  3138

y   1903  GTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCT  1962
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3139  GTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCT  3198

y   1963  GTCTTTTAAAACTGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGCAAGTACAGATA  2022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3199  GTCTTTTTAAAACTGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACACAAGCAAGTACAGATA  3258

y   2023  CCCTTTTCAAGCAAGTGGCAAGTTCAAACAGGATTTTGTGACCAGCGCAGGCGTGGGCCTCCT  2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3259  CCCTTTTCAAGCAAGTGGCAAGTTCAAACAGGATTTTGTGACCAGCGCAGGCGTGGGCCTCCT  3318

y   2083  TCTGCATGATTCATCCAA  2101
          ||||||||||||||||||
b   3319  TCTGCATGATTCTATCCAA  3337
```

Qy   1948 GAGGATCCGTGCCGTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGA 2007
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3139 GAGGATCCGTGCCGTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGA 3198

Qy   2008 AGACAAGTACAGATACCTTTCAAGCAAGTGGCAAGTGCAACAGGATTTTGTGACCAGCG 2067
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   3199 AGACAAGTACAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCG 3258

Qy   2068 CAGGCTGGGCTCCTTCTGCATGATTCATCCAA 2101
           ||||||||||||||||||||||||||||||||
Db   3259 CAGGCTGGGCCTCCTTCTGCATGATTCTATCCAA 3292

RESULT 14
US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US2003017131A1
; GENERAL INFORMATION:
;   APPLICANT: XIAO, XIAO
;   TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;   TITLE OF INVENTION: THEREOF
;   FILE REFERENCE: DE1142
;   CURRENT APPLICATION NUMBER: US/09/845,416
;   CURRENT FILING DATE: 2001-04-30
;   PRIOR APPLICATION NUMBER: 60/200,777
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 3999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-845-416-6

Query Match   76.2%; Score 1602; DB 10; Length 3999;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

Qy   1 GAGCTCATGCCTACACACAGGCTGCTTATGTCACCACCTCCTTCACCACGGAGCCCATT   60
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCCTTCACCACGGAGCCCATT 959

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGGAGAG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGGAGAG 1019

Qy   121 TGAAGTAAACCTGACCGTTATCAAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGACCGTTATCAAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy   241 CCAGTTCATACTTCATGAGGGGTACATGATGGTTTAATGGATCTCCAGAATCAGAAACT 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTCATACTTCATGAGGGGTACATGATGGTTTAATGGATCTCCAGAATCAGAAACT 1199

Qy   301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGCAGGAAAATTATCAGAAGATGAAGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGCAGGAAAATTATCAGAAGATGAAGA 1259

Qy   361 AACTGAAGTACAAGAGCAGATGAATCCTCTAAATTCAAGATGGAATGGCTTCAGGGTAGC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAGATGAATCCTCTAAATTCAAGATGGAATGGCTTCAGGGTAGC 1319

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAAACT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAAACT 1379

Qy   481 GAAAGAGTTGAATGACTGCTAACAAAAACAGGAAGAAGAACAAGGAAATGGAGGAAGA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTTGAATGACTGCTAACAAAAACAGGAAGAAGAACAAGGAAATGGAGGAAGA 1439

959   ATGAGTACTCTTGAGACTGTACGAATATT---------CTGACGAGCAGCCTTTGGA 1008
      ||||||||||||||||||||||||||||| ||||||||||||||||||
2059 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

1009 AGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAA 1068
      ||||||||||| ||||||||||||||||||||||||||||||| ||||| |||||||
2119 CCAGAAGACGGAGCGGAGCATGGCCTGGATAACTTTGCCCGGTGTGGGGATAAATTAGTCCAAAA 2178

1069 TGTCACTCGGCTTCTACGAAGCAGGCAGGCTGAGGAGGTCAATAC---------- 1110
      |||||||||||||||||||||||||||||||||||||||||||||
2179 ACTTGAAAAGAGTACAGCACAGATTTCACAGGCTGTCACCACTCAGCCATCACTAAC 2238

1111 --- 1110
2239 ACAGACAACTGTATGAAAATTGAACCTGCACTCCGTCGACTCGGTGACCAAGAGGGAACAGAGATCCTGGTAAA 2298

1111 ---TGAGTGGGAAAAATTGAACCTGCACTCCGT 1167
         |||||||||||||||||||||||||||||||
2299 GCATGTCAAGAGAGCAACTCCACCACCACCTCCCAAAAGAGAGGCAGATTACTGTGA 2358

1168 CCTTGAAAGACTTCAGGAACTTCAAGAGGCCACGATGAGGCCACGGACCTGGACCTGCGCCA 1227
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2359 TCTTGAAAGACTTCAGGAACTTCAAGAGGCCACGATGAGGCCACGGACCTGGACCTGCGCCA 2418

1228 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGAGTCCTCCA 1287
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2419 AGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCATTGAGTCCTCCA 2478

1288 AGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAAGAGAACGT 1347
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2479 AGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAAGAGAACGT 2538

1348 GAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTA 1407
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2539 GAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTA 2598

1408 TAACCTCCAGCACTCCTGAAGAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGA 1467
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2599 TAACCTCCAGCACTCCTGAAGAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGA 2658

1468 GGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGACTTTGGTCCAGCATCTCAGCACTT 1527
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2659 GGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGACTTTGGTCCAGCATCTCAGCACTT 2718

1528 TCTTTCCACGTCTGTCCACGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTA 1587
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2719 TCTTTCCACGTCTGTCCACGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTA 2778

1588 CTATATCAACCACGAGACTCAAACACTTGCTGGACCATCCCAAAAATGACACGAGCTCTA 1647
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2779 CTATATCAACCACGAGACTCAAACACTTGCTGGACCATCCCAAAAATGACACGAGCCTA 2838

1648 CCAGTCTTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACT 1707
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2839 CCAGTCTTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACT 2898

1708 CCGAAGACTGCAGAAGGCCCTTTGCTTGGATCCTCTTGAGCCTGTCAGTGCATGTGATGC 1767
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2899 CCGAAGACTGCAGAAGGCCCTTTGCTTGGATCCTCTTGAGCCTGTCAGTGCATGTGATGC 2958

1768 CTTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCATGCAATCCTGCGAGATTATTAA 1827
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2959 CTTTGGACCAGCACCAACCTCAAGCAAAATGACCAGCCATGCAATCCTGCGAGATTATTAA 3018

1828 TTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCC 1887
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3019 TTGTTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCC 3078

1888 TCTCTGCGTGGATATGTGTCTGCTGAACTGGCTGCTGGAAGTTTATGATACCGGGACGAACAGG 1947
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3079 TCTCTGCGTGGATATGTGTCTGCTGAACTGGCTGCTGGAAGTTTATGATACCGGGACGAACAGG 3138

```
Qy        541 GCCTCTTGGAACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 1499

Qy        601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACATGGTGGTGTAGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACATGGTGGTGTAGT 1559

Qy        661 TGATGAATCTAGTGAGAATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1619

Qy        721 AGATCGATGGGCAAACATCTGTAGATGACGAAGACCGGCTGGGTTCTTTTACAAGA-- 777
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 AGATCGATGGGCAAACATCTGTAGATGACGAAGACCGGCTGGGTTCTTTTACAAGACCA 1679

Qy        778 -------------------------------------------------------- 777

Db       1680 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCT 1739

Qy        778 -------------------------------------------------------- 777

Db       1740 GGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCCAAACTAGAAATGCCATCTTC 1799

Qy        778 -------------------------------------------------------- 777

Db       1800 CTTGATGTTGGAGGTACCTACTACTACTCATAGATTACTGCAACAGTTCCCCCTGAAAA 1859

Qy        778 -------------------------------------------------------- 777

Db       1860 GTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCG 1919

Qy        778 -------------------------------------------------------- 777

Db       1920 TAAGGAAAAGGCTCCTAGAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGA 1979

Qy        778 -------------------------------------------------------- 777

Db       1980 CCTCCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACCTGGATGAAAAACAGCCA 2039

Qy        778 -------------------------------------------------------- 777

Db       2040 AAAAATCCTGACAGATCCCTGGAAGTTCCGATGATGCAGTCGTTACAAAGACGTTTGGA 2099

Qy        778 -------------------------------------------------------- 777

Db       2100 TAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAAGTCTCTCAACATTAGGTCCCATT 2159

Qy        778 -----CAGTTCTGACCAGGTGGAAGCCGTCTGACCTTTCCTGCAGAACTTCTGGTGTG 831
              |||||||||||||||||||||||||||||||||||||||
Db       2160 GGAAGCCAGTTCTGACCAGGTGGAAGCCGTCTGACCTTTCCTGCAGAACTTCTGGTGTG 2219

Qy        832 GCTACAGCTGAAAGATGATGAATTAAGCCGGCACCGAGGCGCACTTTCCAGC 891
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GCTACAGCTGAAAGATGATGAATTAAGCCGGCACCGAGGCGCACTTTCCAGC 2279

Qy        892 AGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACC 951
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2280 AGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACC 2339

Qy        952 TGTAATCATGAGTACTCTTGACGTGATACGACGAATATTTCTGACAGAGGCCTTTGGAAGG 1011
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2340 TGTAATCATGAGTACTCTTGACGTGATACGACGAATATTTCTGACAGAGGCCTTTGGAAGG 2399

Qy       1012 ACTCGGCTTCTACGAAAGCAGCAGGCTCAATACTGAGTGGGAAAATTGAACCT 1071
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       2400 CACTCGGCTTCTACGAAAGCAGCAGGCTCAATACTGAGTGGGAAAATTGAACCT 2459

Qy       1072 CACTCGGCTTCTACGAAAGCAGCAGGCTCAATACTGAGTGGGAAAATTGAACCT 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2460 CACTCGGCTTCTACGAAAGCAGCAGGCTCAATACTGAGTGGGAAAATTGAACCT 2519
```

```
Qy       1132 GCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCCTGAAAGACTTCCAGGAACTTCA 1191
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2520 GCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCCTGAAAGACTTCCAGGAACTTCA 2579

Qy       1192 AGAGGCCACGATGAGCTGACCTCAAGTGCGCCAAGTGAGGTGATCAAGGGATCCTG 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2580 AGAGGCCACGGATGAGCTGACCTCAAGTGCGCCAAGTGAGGTGATCAAGGGATCCTG 2639

Qy       1252 GCAGCCGTGGCGATCTCCATTGACTTCTCCAAGATCACCTCGAGAAAGTCAAGGC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2640 GCAGCCGTGGCGATCTCCATTGACTTCTCCAAGATCACCTCGAGAAAGTCAAGGC 2699

Qy       1312 ACTTCGAGGAGAAATGCGCCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGCTCG 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2700 ACTTCGAGGAGAAATGCGCCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGCTCG 2759

Qy       1372 CCAGCTTACCACTTGGCATTCAGCTCTTCACCGTATAACCTCAGCACTCTGAAGACCT 1431
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2760 CCAGCTTACCACTTGGCATTCAGCTCTTCACCGTATAACCTCAGCACTCTGAAGACCT 2819

Qy       1432 GAACACCAGATGGAAGCTTCTGCAGGTGGCCATCTCGAGGACCACGAGTCAGGCAGCTGCATGA 1491
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2820 GAACACCAGATGGAAGCTTCTGCAGGTGGCCATCTCGAGGACCACGAGTCAGGCAGCTGCATGA 2879

Qy       1492 AGCCCACAGGACTTTGGTCCAGATCTCAGCACTTCTTCACGTCTGTCCAGGGTCC 1551
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2880 AGCCCACAGGACTTTGGTCCAGATCTCAGCACTTCTTCACGTCTGTCCAGGGTCC 2939

Qy       1552 CTGGGAGAGAGCCATCTCGCAAACAAAGTGCCCTACTATCAACCAGAGACTCAAAC 1611
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2940 CTGGGAGAGAGCCATCTCGCAAACAAAGTGCCCTACTATCAACCAGAGACTCAAAC 2999

Qy       1612 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGCTGACCTGAATAA 1671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3000 AACTTGCTGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGCTGACCTGAATAA 3059

Qy       1672 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAACTGCAGAAGGCCCTTTG 1731
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3060 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAACTGCAGAAGGCCCTTTG 3119

Qy       1732 CTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACACAACCTCAAGCA 1791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3120 CTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACACAACCTCAAGCA 3179

Qy       1792 AAATGACCAGCCCATGGATATCCTGCGAGATTAATTAATTGTTGACCACTATTATGACCG 1851
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3180 AAATGACCAGCCCATGGATATCCTGCGAGATTAATTAATTGTTGACCACTATTATGACCG 3239

Qy       1852 CCTGGAGCAAGAGCACAACATTTGGTCAACGTCCCTCTTCGGCGTGGATAATGTGTCTGAA 1911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3240 CCTGGAGCAAGAGCACAACATTTGGTCAACGTCCCTCTTCGGCGTGGATAATGTGTCTGAA 3299

Qy       1912 CTTGGCTGCTGAATGTTATGATACGGGACAGAGGGGATCCGTGTCCTGTCTTTAA 1971
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 CTTGGCTGCTGAATGTTATGATACGGGACAGAGGGGATCCGTGTCCTGTCTTTAA 3359

Qy       1972 AACTGGCATCATTCCCGTGTAAAGCACATTGGAAGACAAGCAAGTACAGATACCTTTCAA 2031
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 AACTGGCATCATTCCCGTGTAAAGCACATTGGAAGACAAGCAAGTACAGATACCTTTCAA 3419

Qy       2032 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCCTGGGCCTTCCTCCTGCATGA 2091
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCCTGGGCCTTCCTCCTGCATGA 3479

Qy       2092 TTCTATCCAA 2101
              ||||||||||
Db       3480 TTCTATCCAA 3489
```

RESULT 15
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US2003017131 2A1

GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 28
LENGTH: 4966
TYPE: DNA
ORGANISM: Homo sapiens
3-09-845-416-28

Query Match   76.2%;   Score 1602;   DB 10;   Length 4966;
Best Local Similarity   81.1%;   Pred. No.0;
Matches 2101;   Conservative   0;   Mismatches   0;   Indels   489;   Gaps   1;

```
        1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCGACCCTTACACCTTACACGGAGCCCATT   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTCGACCCTACACGGAGCCCATT     1716

       61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG    120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG    1776

      121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTTCTTTC    180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTTCTTTC    1836

      181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA    240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA    1896

      241 CCAGTTTCATACTTCATGATGAGGGGTACATCATGATGAGGGTACAGCCCATCACAGGGCTTGG    300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1897 CCAGTTTCATACTTCATGATGAGGGGTACATCATGATGAGGGTACAGCCCATCACAGGGCTTGG    1956

      301 TAATATTCTACAATTGGGAAGTAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA    360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1957 TAATATTCTACAATTGGGAAGTAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA    2016

      361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTGAGC    420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTGAGC    2076

      421 TAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAATGCATAGAGTTTAATGGATCTCCAGAAATCAGAAACT    480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     2077 TAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAATGCATAGAGTTTAATGGATCTCCAGAAATCAGAAACT    2136

      481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAATGGAGAAGAA    540
          |||||||||||||||||||||||||||||||||||||||||||||||||
     2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAATGGAGAAGAA    2196

      541 GCCTCTTCGTGGACCTGACTCTTGAAGACTCTGAAGTCTTGAAGGGTGCTTCA    600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
     2197 GCCTCTTCGTGGACCTGACTCTTGAAGACTCTGAAGTCTTGAAGGGTGCTTCA    2256

      601 AGAAGATCTAGAACAAGAACAAGAGTCAGGGTCAATTCTCCACTCACTTCACATGGTGGTGGTAGT    660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     2257 AGAAGATCTAGAACAAGAACAAGAGTCAGGGTCAATTCTCCACTCACTTCACATGGTGGTGGTAGT    2316

      661 TGATGAATCTAGTGGAAGAACAACTTAAGGTATTGGG    720
          |||||||||||||||||||||||||||||||||||||
     2317 TGATGAATCTAGTGGAAGAACAACTTAAGGTATTGGG    2376

      721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGGGTTCTTTTACAAGA---    777
          ||||||||||||||||||||||||||||||||||||||||||||||  ||||
     2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCGGGTTCTTTTACAAGACCA    2436
```

```
Qy  777 -------------------------------------------------------------------    777
Db 2437 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCCTACTGTTACTCT    2496
Qy  778 -------------------------------------------------------------------    777
Db 2497 GGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC    2556
Qy  777 -------------------------------------------------------------------    777
Db 2557 CTTGATGTTGGGAGGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAA    2616
Qy  778 -------------------------------------------------------------------    777
Db 2617 GTTTCTTGCTTGCTTACGAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCG    2676
Qy  778 -------------------------------------------------------------------    777
Db 2677 TAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAGAGCTGATGAAACATGGCAAGA    2736
Qy  777 -------------------------------------------------------------------    777
Db 2737 CCTCCAAGGTGAAATTGAAGCTCACCAGATGTTTATCACAACCTGATGAAAACAGCCA    2796
Qy  777 -------------------------------------------------------------------    777
Db 2797 AAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTTCCTGTTACAAAGACGTTTGGA    2856
Qy  777 -------------------------------------------------------------------    777
Db 2857 TAACATGAGAACTTCAAGTGGAGTGAACTTCGGAGAAAAAGTCTCTCAACATTAGTCCCATTT    2916
Qy  778 -------------------------------------------------------------------    831
Db 2917 GGAAGCCAGTTCTGACCAGTGGAAGCCGTCTGCACCTTTTCTCGCAGGAACTTCTCGGTGTG    2976
Qy  832 GCTACAGCTGAAAGATGAATGATGAATTAAGCCGGCACACCTATTGGAGGGACTTCCAGC    891
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 2977 GCTACAGCTGAAAGATGAATGATGAATTAAGCCGGCACACCTATTGGAGGGACTTCCAGC    3036
Qy  892 AGTTCAGAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGAACC    951
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3037 AGTTCAGAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGAACC    3096
Qy  952 TGTAATCATGAGTACTCTTGAGACTGTACAGAGACATTCTGACAGAATATTCTGACAGATT    1011
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3097 TGTAATCATGAGTACTCTTGAGACTGTACAGAGACATTCTGACAGAATATTCTGACAGATT    3156
Qy 1012 ACTAGAGAAACTCCACCAGGAGCCCCAGAGAGCTTGCCTCCTGAGGAATCAGAGAATGT    1071
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 3157 ACTAGAGAAACTCCACCAGGAGCCCCAGAGAGCTTGCCTCCTGAGGAATCAGAGAATGT    3216
Qy 1072 CACTCGGCTTCTACGAAGCAGCAGGAGGTCAATACTGACTCAATACTGAGTGGGAAAATTGAACCT    1131
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 3217 CACTCGGCTTCTACGAAAGCAGCAGGAGGTCAATACTGACTCAATACTGAGTGGGAAAATTGAACCT    3276
Qy 1132 GCAGTCCGCTGACTGGCACGGAGAAAAATAGAGATGAGACCCTGAAAGACTCCAGGAACTTCA    1191
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3277 GCAGTCCGCTGACTGGCACGGAGAAAAATAGAGATGAGACCCTGAAAGACTCCAGGAACTTCA    3336
Qy 1192 AGAGGCCACGGATGAGCTGGACCTGAGTCGGCCAAGCTGAGGTGATCAAGGGATCCTG    1251
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 3337 AGAGGCCACGGATGAGCTGGACCTGAGTCGGCCAAGCTGAGGTGATCAAGGGATCCTG    3396
Qy 1252 GCAGCCCGTGGCGATCTCTCATTGACTCTTCCAAGATCACCTCGAGAAAGTCAAGGC    1311
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3397 GCAGCCCGTGGCGATCTCTCATTGACTTCCTTCCAAGATCACCTCGAGAAAGTCAAGGC    3456
Qy 1312 ACTTCGAGGAGAAAATTGCGCCTCTGAAAGAGACGTGAGCCACTTGCTCG    1371
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3457 ACTTCGAGGAGAAAATTGCGCCTCTGAAAGAGACGTGAGCCACTTGCTCG    3516
Qy 1372 CCAGCTTACCACTTTGGCATTCAGCTCTCACCGTATAACCTCAGACACTCAGAAGACCT    1431
```

```
b            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      3517   CCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCACTCTGAAGACCT    3576

y            GAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGAGGACCAGGCAGCTGCATGA         1491
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b            GAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGAGGACCAGGCAGCTGCATGA         3636
      3577                                                                  3636

y            AGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGGTCC    1551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      3637  AGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGGTCC    3696

y      1552  CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC  1611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      3697  CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC  3756

y      1612  AACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA  1671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      3757  AACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA  3816

y      1672  TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG  1731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      3817  TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTG  3876

y      1732  CTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACCTATTTATGACCG  1791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      3877  CTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACCTATTTATGACCG  3936

y      1792  AAATGACCAGCCAGCCCATGGATATCCTCGCAGATTATTAATTGTTTGACCACTATTTATGACCG  1851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      3937  AAATGACCAGCCAGCCCATGGATATCCTCGCAGATTATTAATTGTTTGACCACTATTTATGACCG  3996

y      1852  CCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGATAATGTGTCTGAA  1911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      3997  CCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGATAATGTGTCTGAA  4056

y      1912  CTGGCTGCTGAATGTTTATGATACCGGACCGAACAGGGAGGATCCGTGTCCTGTCTTTTAA  1971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      4057  CTGGCTGCTGAATGTTTATGATACCGGACCGAACAGGGAGGATCCGTGTCCTGTCTTTTAA  4116

y      1972  AACTGGCATCATTTCCCTGTGTAAAGCACATTTGAAGACAAGTACAGATACCTTTTCAA  2031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      4117  AACTGGCATCATTTCCCTGTGTAAAGCACATTTGAAGACAAGTACAGATACCTTTTCAA  4176

y      2032  GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGCGAGGCCTCCTTCTGCATGA       2091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      4177  GCAAGTGGCAAGTTCAACAGGATTTTGTGACCACGCGAGGCCTGGGCCTCCTTCTGCATGA  4236

y      2092  TTCTATCCAA                                                     2101
             ||||||||||
b      4237  TTCTATCCAA                                                     4246
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on: April 4, 2004, 11:55:32 ; Search time 542.141 Seconds
(without alignments)
16463.377 Million cell updates/sec

tle: US-09-845-416-12_COPY_900_3000
rfect score: 2101
quence: 1 gagctatgctacacacagg........ttctgcgatgattctatccaa 2101

coring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

arched: 3373863 seqs, 2124099041 residues

tal number of hits satisfying chosen parameters: 6747726

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Length	DB	ID	Description
1	2101	100.0	3510	6	AAD37240	Aad37240 Human dys
2	2101	100.0	4476	6	AAD37259	Aad37259 Adeno-ass
3	1760.4	83.8	5339	6	ABK81998	Abk81998 DNA encod
4	1750.8	83.3	3531	6	AAD37238	Aad37238 Human dys
5	1750.8	83.3	4498	6	AAD37258	Aad37258 Adeno-ass
6	1743	83.0	3858	6	AAD37237	Aad37237 Human dys
7	1743	83.0	4825	6	AAD37257	Aad37257 Adeno-ass
8	1743	83.0	4848	6	AAD37263	Aad37263 Adeno-ass
9	1743	83.0	5060	6	AAD37264	Aad37264 Adeno-ass
10	1641	78.1	4414	6	AAD37260	Aad37260 Adeno-ass
11	1629	77.5	3446	6	AAD37242	Aad37242 Human dys
12	1629	77.5	5462	6	ABK81999	Abk81999 DNA encod
13	1611.6	76.7	5417	6	ABK81997	Abk81997 DNA encod
14	1602	76.2	3999	6	AAD37234	Aad37234 Human dys
15	1602	76.2	4966	6	AAD37256	Aad37256 Adeno-ass
16	1602	76.2	4990	6	AAD37262	Aad37262 Adeno-ass
17	1419	67.5	4182	6	AAD37230	Aad37230 Human dys
18	1419	67.5	5149	6	AAD37255	Aad37255 Adeno-ass
19	1328.8	63.2	2169	6	AAD37232	Aad37232 Human dys
20	1328.8	63.2	5952	5	AAD06794	Aad06794 Human dys
21	1328.8	63.2	8689	6	ABK82000	Abk82000 DNA encod
22	1328.8	63.2	11058	6	AAD37229	Aad37229 Human dys
23	1328.8	63.2	11241	6	ABK82005	Abk82005 cDNA enco
24	1328.8	63.2	11443	6	ABK82002	Abk82002 DNA encod
25	1328.8	63.2	12923	1	AAN90338	Aan90338 Sequence
26	1328.8	63.2	13957	6	ABK81959	Abk81959 cDNA enco
27	1328.8	63.2	13957	6	ABT10904	Abt10904 Human bre
28	1328.8	63.2	13957	6	ABN95786	Abn95786 Gene #228
29	1328.8	63.2	13957	6	ABS69900	Abs69900 Human dys
30	1324	63.0	1821	6	AAD37241	Aad37241 Human dys
31	1317.8	62.7	13977	6	ABS70403	Abs70403 Human bon
32	1317	62.7	4402	3	AAZ48567	Aaz48567 A rod sho
33	1196.6	57.0	4402	3	AAZ48568	Aaz48568 A rod sho
34	1149.6	54.7	13815	6	ABK81960	Abk81960 cDNA enco
35	1149.6	54.7	13815	6	ABI99799	Abi99799 Mouse isc
36	1149.6	54.7	19307	2	AAT27558	Aat27558 Shuttle v
c 37	1148	54.6	13815	2	AAV18885	Aav18885 Mus muscu
38	937	44.6	1434	6	AAD37243	Aad37243 Human dys
39	785.4	37.4	1991	6	AAD37231	Aad37231 Human dys
40	777	37.0	1667	6	AAD37235	Aad37235 Human dys
41	718.8	34.2	4075	3	AAZ48569	Aaz48569 A rod sho
42	667.2	31.8	3747	3	AAZ48566	Aaz48566 A rod sho
43	618.6	29.4	3275	1	AAN97129	Aan97129 Partial s
44	613.6	29.2	3163	3	AAZ48571	Aaz48571 A rod sho
45	555.6	26.4	10705	7	ABT41896	Abt41896 Toxicity

ALIGNMENTS

RESULT 1
AAD37240
AAD37240
ID   AAD37240 standard; DNA; 3510 BP.
XX
AC   AAD37240;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3510.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 51-52; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

```
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CX   CR domain) and 11047-11058 (dystrophin last 3 amino acids)
SQ   Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;

Query Match             100.0%;  Score 2101;  DB 6;  Length 3510;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 2101;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy       1 GAGCTATGCCACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 GAGCTATGCCACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 959

Qy      61 TCCTTCACAGCAGTTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     960 TCCTTCACAGCAGTTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1019

Qy     121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGAGTATTATCGTGGCTTCTTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGAGTATTATCGTGGCTTCTTTC 1079

Qy     181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy     241 CCAGTTCATCATGATGATGATACAGCCCATCAGGGCCGGGTTGG 300
           ||||||||||||||||||||||||||||||||||||||||||||
Db    1140 CCAGTTCATCATGATGATGATACAGCCCATCAGGGCCGGGTTGG 1199

Qy     301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAAT TATCAGAAGATGAAGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1259

Qy     361 AACTGAAGTACAAGAGCAGCAGATGAATCCTCAAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1260 AACTGAAGTACAAGAGCAGCAGATGAATCCTCAAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1319

Qy     421 TAGCATGAAAACAAAGCAATTACAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1320 TAGCATGAAAACAAAGCAATTACAATTTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 1379

Qy     481 GAAAGAGTTGATGACTGGCTAACAAAAACAGAGAACAAGGAAAATGAGGAGAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1380 GAAAGAGTTGATGACTGGCTAACAAAAACAGAGAACAAGGAAAATGAGGAGAA 1439

Qy     541 GCCTCTTGGACCTGATCTTGAAGACCTGATCTTGAAGACGTACAACACATAAGGTGCTTCA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1440 GCCTCTTGGACCTGATCTTGAAGACCTGATCTTGAAGACGTACAACACATAAGGTGCTTCA 1499

Qy     601 AGAAGATCTAGAACAAGAGAGTCAGGGTCAATTCTCTCACTCCAGACAACCGCGGTGGGGTTCTTTTACAAGACAG 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1500 AGAAGATCTAGAACAAGAGAGTCAGGGTCAATTCTCTCACTCCAGACAACCGCGGTGGGGTTCTTTTACAAGACAG 1559

Qy     661 TGATGAATCTAGTGAGGAAATCTGTAGATGGAGACAAGAACCGCGCTGGGGTCTTTTCTTTTACAAGACAG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1560 TGATGAATCTAGTGAGGAAATCTGTAGATGGAGACAAGAACCGCGCTGGGGTCTTTTCTTTTACAAGACAG 1619

Qy     721 AGATCGATGGGCAAATCTGTAGATGGGCAAATCTGTAGATGGAGACGCGCTGGGGTCTTTTCTTTTACAAGACAG 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1620 AGATCGATGGGCAAATCTGTAGATGGGCAAATCTGTAGATGGAGACGCGCTGGGGTCTTTTCTTTTACAAGACAG 1679

Qy     781 TTCTGACCAGTGGAAGCTGGCTCTGCACCTTTCTGGTGGTTCTGGTGGTGGCTACAGCT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1680 TTCTGACCAGTGGAAGCTGGCTCTGCACCTTTCTGGTGGTTCTGGTGGTGGCTACAGCT 1739

Qy     841 GAAAGATGAATGAATTAAGCCGGCAGCCGGCGACTTTCCAGCAGTTCAGAA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1740 GAAAGATGAATGAATTAAGCCGGCAGCCGGCGACTTTCCAGCAGTTCAGAA 1799

Qy     901 GCAGAACGACGATGTACATAGGGCCTCAAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCAT 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1800 GCAGAACGACGATGTACATAGGGCCTCAAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCAT 1859
```

```
Qy     961 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCCTTGGAAGGACTAGAGAA 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1860 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCCTTGGAAGGACTAGAGAA 1919

Qy    1021 ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTCGGAGGAGAGCCCAGAATGTCACTCGGCT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1920 ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTCGGAGGAGAGCCCAGAATGTCACTCGGCT 1979

Qy    1081 TCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTCCGC 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1980 TCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTCCGC 2039

Qy    1141 TGACTGCAGAGAAAATAGATGAGACCCTTGAAAGACTTCAAGAACTTCAAGAGGCCAC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2040 TGACTGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAACTTCAAGAGGCCAC 2099

Qy    1201 GGATGAGCTGGACCTCAAGCTGCGCCAGGCTGAGTGTGATCAAGGGATCCTGCGAGCCCGT 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2100 GGATGAGCTGGACCTCAAGCTGCGCCAGGCTGAGTGTGATCAAGGGATCCTGCGAGCCCGT 2159

Qy    1261 GGGCGATCCTCCTCATTGACTTCTCTCCAAGATCACCTCGAGAAGTCACGGCACTTCGAGG 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2160 GGGCGATCCTCCTCATTGACTTCTCTCCAAGATCACCTCGAGAAAGTCACGGCACTTCGAGG 2219

Qy    1321 AGAAATTGCGCTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCGTTAC 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2220 AGAAATTGCGCTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCGTTAC 2279

Qy    1381 CACTTTGGCATTCAGTCTCAGCTGTGAGGACCGAGTCAGCAGTCTGCATGAAGCCACAG 1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2280 CACTTTGGCATTCAGTCTCAGCTGTGAGGACCGAGTCAGCAGTCTGCATGAAGCCACAG 2339

Qy    1441 ATGGAAGCTTCTGCAGGTGGCCGTGCGAGTCGCACTTCTTCTGCCACGTCGTCCAGGGTCCTGGGAGAG 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2340 ATGGAAGCTTCTGCAGGTGGCCGTGCGAGTCGCACTTCTTCTGCCACGTCGTCCAGGGTCCTGGGAGAG 2399

Qy    1501 GGACTTGGTGGTCCAGATCTCCAGCACTTCTTCTTCCAGGGTCCTGTCCCGTGGGAGAG 1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2400 GGACTTGGTGGTCCAGATCTCCAGCACTTCTTCTTCCAGGGTCCTGTCCCGTGGGAGAG 2459

Qy    1561 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACAAGAACTCAAACAACTTGCTG 1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2460 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACAAGAACTCAAACAACTTGCTG 2519

Qy    1621 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACGTGACCGAATAATGTCAGATT 1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2520 GGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACGTGACCGAATAATGTCAGATT 2579

Qy    1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCGAGAAGGCGCCCTTTGCTTGGTAGT 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2580 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCCGAAGAAGGCGCCCTTTGCTTGGTAGT 2639

Qy    1741 CTTGAGCCTGTCAGTGCATGTGATGCCTGCGTCATGTGATGCCTGTCCTGTCTCTTTAAAAACTGGCAT 1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2640 CTTGAGCCTGTCAGTGCATGTGATGCCTGCGTCATGTGATGCCTGTCCTGTCTCTTTAAAAACTGGCAT 2699

Qy    1801 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA 1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2700 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGCCTGGAGCA 2759

Qy    1861 AGAGCACAACATTTTGGTCAACGTCCCTCTCTGCGTGATATGTGTCTGAACTCGCTGCT 1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2760 AGAGCACAACATTTTGGTCAACGTCCCTCTCTGCGTGGTGATATGTGTCTGAACTCGCTGCT 2819

Qy    1921 GAATGTTTATGATCACGGGAACAGGGAGGAGGATCCGTCCTGTCCGTCTGTCTCTTTAAAAACTGGCAT 1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2820 GAATGTTTATGATCACGGGAACAGGGAGGAGGATCCGTCCTGTCCGTCTGTCTCTTTAAAAACTGGCAT 2879

Qy    1981 CATTTCCCTGTGTAAAGCACATTTGAAGAGACAAGAAGCAAGTACAGATACCTTTCAAGCAAGTGGC 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2880 CATTTCCCTGTGTAAAGCACATTTGAAGAGACAAGAAGCAAGTACAGATACCTTTCAAGCAAGTGGC 2939
```

AD37259
AAD37259   standard; DNA; 4476 BP.
AAD37259;
21-AUG-2002 (first entry)
Adeno-associated virus vector plasmid, AAV-MCK-3510.
Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin; adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD; Becker muscular dystrophy; ds.
Homo sapiens.
Unidentified.
Chimeric.
WO200183695-A2.
08-NOV-2001.
27-APR-2001; 2001WO-US013677.
28-APR-2000; 2000US-0200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular dystrophy minigene. The minigene comprises an N-terminal domain or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence.
Example 1; Page 63-65; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence.

Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;
Query Match   100.0%; Score 2101; DB 6; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTCGACCCTACACGGAGCCCATT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1656 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTCGACCCTACACGGAGCCCATT 1715

Qy   61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DT   13-AUG-2002 (first entry)
DE   DNA encoding mini-dystrophin protein deltaR2-R21.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens. .
OS   Synthetic.
XX
PN   WO200229056-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Example 6; Fig 13; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 U; 0 Other;

Query Match   83.8%; Score 1760.4; DB 6; Length 5339;
Best Local Similarity   90.3%; Pred. No. 0;
Matches 1910; Conservative   0; Mismatches 191; Indels 15; Gaps   2;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACGACATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAG 120
Db   1159 TCCTTCACGACATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGGAG 1218

Qy   121 TGAAGTAAACTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
Db   1219 TGAAGTAAACTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGCAGGAGGATTCTAATGATGTGGAAGTGGTGAAAGA 240
Db   1279 TGCTGAGGACACATTGCAAGCACAAGCAGGAGGATTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTTCATACTCATGAGGAAGTAAGCTGATTGAACAGCCCATCAGGGGCCGGGTTGG 300
Db   1339 CCAGTTTCATACTCATGAGGAAGTAAGCTGATTGAACAGCCCATCAGGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGAACAGGAAAAATTATCAGAAGATGAAGA 360
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGAACAGGAAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGACAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420

Db   2796 TGACTGGCAGAGAAATAGATGAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGCCAC 2855
Qy   1201 GGATGAGACTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGCAGCCCGT 1260
Db   2856 GGATGAGACTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGCAGCCCGT 2915
Qy   1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 1320
Db   2916 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 2975
Qy   1321 AGAAATTGCGCCTCTGAAAAGAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 1380
Db   2976 AGAAATTGCGCCTCTGAAAAGAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 3035
Qy   1381 CACTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGACACCAG 1440
Db   3036 CACTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGACACCAG 3095
Qy   1441 ATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG 1500
Db   3096 ATGGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAG 3155
Qy   1501 GGACTTTGGTCCAGCATCTCAGCACTTCTTCTTCCACGTCGTGTCCAGGGTCCCTGGGAGAG 1560
Db   3156 GGACTTTGGTCCAGCATCTCAGCACTTCTTCTTCCACGTCGTGTCCAGGGTCCCTGGGAGAG 3215
Qy   1561 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 1620
Db   3216 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 3275
Qy   1621 GGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 1680
Db   3276 GGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 3335
Qy   1681 CTCAGCTTATAGGACTGCCATGAAAACTCCGAGAGACTGCAGAAGGCCCCTTTGCTTGGATCT 1740
Db   3336 CTCAGCTTATAGGACTGCCATGAAAACTCCGAGAGACTGCAGAAGGCCCCTTTGCTTGGATCT 3395
Qy   1741 CTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACACAACCTCAAGCAAAATGACCA 1800
Db   3396 CTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACACAACCTCAAGCAAAATGACCA 3455
Qy   1801 GCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTATTGATGTTGACCGCCTGGAGCA 1860
Db   3456 GCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTATTGATGTTGACCGCCTGGAGCA 3515
Qy   1861 AGAGCACACAACAATTGGTCAAGTCCCTCTCCGTGGATATGTGTCGAACTGGCTGCT 1920
Db   3516 AGAGCACACAACAATTGGTCAAGTCCCTCTCCGTGGATATGTGTCGAACTGGCTGCT 3575
Qy   1921 GAATGTTTATGATACGGGACGAACAGGGAGATCCGTGTCCTGTGTCTTTAAAACTGGCAT 1980
Db   3576 GAATGTTTATGATACGGGACGAACAGGGAGATCCGTGTCCTGTGTCTTTAAAACTGGCAT 3635
Qy   1981 CATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAGACCTTTTCAAGCAAGTGGC 2040
Db   3636 CATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAGACCTTTTCAAGCAAGTGGC 3695
Qy   2041 AAGTTCAACAGGATTTTGTGACCAGCGCAGCCGGCTGGGCCTTCCTCTTCTGCATGATTCTATCCA 2100
Db   3696 AAGTTCAACAGGATTTTGTGACCAGCGCAGCCGGCTGGGCCTTCCTCTTCTGCATGATTCTATCCA 3755

Qy   2101 A 2101
Db   3756 A 3756

RESULT 3
ABK81998 standard; DNA; 5339 BP.
ID   ABK81998
KX
AC   ABK81998;
XX

```
1459 AACTGAAGTACAAGAGCAGAATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

  421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
       ||||||||||||||||||||||||||||| || |||||||||||||||||||||
 1519 TAGCATGGAAAAACAAAGCAATTTACATCAGATTACTGCAACAGTTCCCCTGGACCT 1578

  481 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAACAGGAAAATGAGAGGAAGA 540
       |||||| || ||||| ||||||  ||| | || || ||  || | || |||||  |
 1579 GGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCCAATGTCCTACAGGATGC 1638

  541 GCCTCTTGGACCTGATCTTGAAGACTCCAAAACGCCAAGTACAACAACATAAGGTGCTTCA 600
                                  |||     |   |  |  ||||
 1639 TACCCGTAAGGAAAGGCTCCAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATG 1698

  601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCTCACTCACATGGTGGTGGTAGT 660
       || |  |||  |||||| | |  || ||  ||||  ||  |||| |  |  |||
 1699 GCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCCAACCTGGATGAAAA 1758

  661 TGATGA-------ATTAGTGGATGATCACCCAACTGCTCGTTTGGAAGACAACACT 708
                     | ||     |   | || |||| |||   |||  |||||||||
 1759 CAGCCAAAAAATCCTGAGATCCCTGAAGGTTCCGATGATGCAGTCCGTTACAAAGACG 1818

  709 TAAGG---TATTGGGAGGATGATGGGCAAAACATCTGTAGAATGGACAGAAGACCGCTGGGT 765
             |  | ||    ||||  | ||||||| |  |   ||| | ||||| | |  |
 1819 TTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAGAAAAGTCTCTCAACATTAGGTC 1878

  766 TCTTTTACAAGACAGTTCTGACCAGTGGAAGCGTCGACCTTTCTGCAGGAACTTCT 825
       |||| ||||||||||||||||||||| |||  ||  |  ||||||||||||| |
 1879 CCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGACCTTTCTGCAGGAACTTCT 1938

  826 GGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTT 885
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1939 GGTGTGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTT 1998

  886 TCCAGCAGTTCAGAAGCAGAACGAGATGACTAGGCCTTCAAGAGGGAATTGAAAACTAA 945
       |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
 1999 TCCAGCAGTTCAGAAGCAGAACGAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAA 2058

  946 AGAACCTGTAATCATGAGTACTCTTCTACCAGGAGCGCCCAGAAGCAGGCTGAGTGGGAAAATT 1005
       |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
 2059 AGAACCTGTAATCATGAGTACTCTTCTACCAGGAGCGCCCAGAAGCAGGCTGAGTGGGAAAATT 2118

 1006 GGAAGGACTAGAGAAAACTCTACCAGGAGCGCCTGCCTCCGAGGAGGAGGAGGCCCA 1065
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2119 GGAAGGACTAGAGAAAACTCTACCAGGAGCGCCTGCCTCCGAGGAGGAGGAGGCCCA 2178

 1066 GAATGTCACTCGCCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATT 1125
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2179 GAATGTCACTCGCCTTCTACGAAAGCAGGCTCAATACTGAGTGGGAAAAATT 2238

 1126 GAACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCCTTGAAAGACTCCAGGA 1185
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2239 GAACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCCTTGAAAGACTCCAGGA 2298

 1186 ACTTCAAGAGGCCACGGATGAGCTGGACCTGAAGCTGCGCCAAGCTGAGGTGATCAAGGG 1245
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2299 ACTTCAAGAGGCCACGGATGAGCTGGACCTGAAGCTGCGCCAAGCTGAGGTGATCAAGGG 2358

 1246 ATCCTGCAGCCGCGTGGGCAGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAGT 1305
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2359 ATCCTGCAGCCGCGTGGGCAGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAGT 2418

 1306 CAAGGCACTTCGAGGAGAAAATTGCGCCTGAAAGAAGACGTGAGCCACGTCAATGACCT 1365
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2419 CAAGGCACTTCGAGGAGAAAATTGCGCCTCCTGAAAGAGAACGTGAGCCACGTCAATGACCT 2478

 1366 TGCTCGCCAGCTTACCACTTTGGGCCATTCAGCTCCAGCATCACCGTATAACCTTCCAGCACTCTGGA 1425
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2479 TGCTCGCCAGCTTACCACTTTGGGCCATTCAGCTCCAGCATATAACCTTCCAGCACTCTGGA 2538

 1426 AGACCTGAACACCAGGATGAGCTGGAAGCTTCTGCAGGTGCCGTCGAGGACCGAGTCAGGCAGCT 1485
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2539 AGACCTGAACACCAGGATGAGCTGGAAGCTTCTGCAGGTGCCGTCGAGGACCGAGTCAGGCAGCT 2598
```

```
Qy 1486 GCATGAAGGCCACAGGGACTTTGGTCAGCATCCAGCACTTCTTCCACGTCTGTCCA 1545
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2599 GCATGAAGGCCCACAGGGACTTTGGTCAGCATCCAGCACTTCTTCCACGTCTGTCCA 2658

Qy 1546 GGGTCCCTGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGAC 1605
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2659 GGGTCCCTGGGAGAGAGCCATCTCGCCAAACACAAGTGCCCTACTATATCAACCACGAGAC 2718

Qy 1606 TCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCT 1665
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2719 TCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCT 2778

Qy 1666 GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGGAAACTCCGAAGACTGCAGAAGGC 1725
        ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2779 GAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGGAAACTCCGAAGACTGCAGAAGGC 2838

Qy 1726 CCTTTGCTTGGATCTCTTGAGCGTCGTCAGCTGCATGCATGCCTTGGACCAGCACCAACCT 1785
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2839 CCTTTGCTTGGATCTCTTGAGCCGTCGTCAGCTGCATGCATGCCTTGGACCAGCACCAACCT 2898

Qy 1786 CAAGCAAAATGACCAGCCCATGGATCCTGCAGATTATTAATTGTTTGACCACTATTTA 1845
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2899 CAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTTA 2958

Qy 1846 TGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCGCGTGGATATGTG 1905
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2959 TGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCGCGTGGATATGTG 3018

Qy 1906 TCTGAACTGGCTGCTGAATGTTTATGATACGGGAGCAACGAGGAGATCCGTGTCCGTC 1965
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3019 TCTGAACTGGCTGCTGAATGTTTATGATACGGGAGCAACGAGGAGATCCGTGTCCGTC 3078

Qy 1966 TTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAAGATACCT 2025
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3079 TTTTAAAACTGGCATCATTTCCCTGTGTGTAAAGCACACATTTGGAAGACAAGTACAAGATACCT 3138

Qy 2026 TTTCAAGCAAGTGGCAAGTTCAACAGATTTGTGACCAGCGCAGCTGGCCTCCTTCT 2085
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3139 TTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGCTGGCCTCCTTCT 3198

Qy 2086 GCATGATTCTATCCAA 2101
        ||||||||||||||||
Db 3199 GCATGATTCTATCCAA 3214
```

```
RESULT 4
AAD37238
ID   AAD37238   standard; DNA; 3531 BP.
XX
AC   AAD37238;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Human dystrophin minigene delta3531.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
```

X
IR   WPI; 2002-049342/06.
X
T   New dystrophin minigene for treating Duchenne or Becker muscular
T   dystrophy comprises an N-terminal domain or modified N-terminal domain,
T   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
T   gene.
X
S   Example 1; Page 50-51; 71pp; English.
X
C   The present invention relates to an isolated nucleotide sequence encoding
C   a dystrophin minigene. The minigene comprises N-terminal or modified N-
C   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
C   domains and cysteine-rich domains of dystrophin or utrophin genes. The
C   invention also relates to a recombinant adeno-associated virus (AAV)
C   comprising dystrophin minigene operably linked to an expression control
C   element. The dystrophin minigene in operable linkage with an expression
C   control element, in a recombinant adeno-associated virus or retrovirus is
C   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
C   dystrophy (BMD) in a mammalian subject. The present sequence is human
C   dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
C   hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR
C   domain) and 11047-11058 (dystrophin last 3 amino acids)
X
Q   Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;

Query Match   63.3%;   Score 1750.8;   DB 6;   Length 3531;
Best Local Similarity   90.2%;   Pred. No. 0;
Matches 1914;   Conservative   0;   Mismatches   187;   Indels   21;   Gaps   3;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTGACCCTACACGAGCCCATT   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT   959

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGCAAGTCATTGGCAGTCATTGATGGGAGAG   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTTGGAAGCTCCTGAAGCAAGTCATTGGCAGTTCATTGATGGGAGAG 1019

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTATTATCTGGCTTCTTTC   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTATTATCTGGCTTCTTTC   1079

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy   241 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGCACAGCCCATCAGGGCCGGGTTGG   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG 1199

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259

Qy   361 AACTGAAGTACAAGAGACAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGACAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC 1319

Qy   421 TAGCATGGAAAAACAAAAGCAATTTACATAGAGTT------TTAATGGATCTCCAGAATCA   474
           |||||||||||||||||||||| |||||||||||||
Db   1320 TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTACTAGAGATTACTACTGCAACAGTTCCCCCT 1379

Qy   475 GAAACTGAAAGAGTTGAATGACTGGCTACAAAAACAGAAGAACAAGGAAAATGGA   534
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GGACCTGGAAAAGTTTCTTGCCTGGCTTACAAAACAGAAGAACAACTGCCAATGCCTTACA 1439

Qy   535 GGAAGAGCCTCCTTGGACCTGATCTTGAAGAACCTAAAACGCCAAGTACAACAACATAAGGT   594
           ||| ||||||||||||||||||
Db   1440 GGATGCTACCCGTAAGGAAGGGTCCTAGAAGAGCTCCTAGAAGAGAGCTGAATGATGGTAA 1499

Qy   595 GCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACCATGTGGT   654
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 ACAATGGCAAGACCTCCAAGGTGAAATTGAAGTCAACAGATGTTTATCAGAACCTGGA 1559

Qy   655 GGTAG-------------------------TTGATGAATCTAGTGAGATCACGCAACTGTGCTTTGGAAGA   702
               | |||||||||||||||||||||
Db   1560 TGAAAACAGGCCAAAAAAATCCTGAGATCCCTGGAGGTTCCGATGGGAAGGTCCGTGTTACA 1619

Qy   703 ACAACTTAAGG---TATTGGGAGATCGATGGGCCAAACATCTGTGATGGACAGAGACCG   759
           |||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   1620 AAGACGTTTGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTCTCAACAT 1679

Qy   760 CTGGGTTCTTTTACAAGACAGTTCTGACCAGTGGAAGGGTCTGCCACCTTTCTCTGCGACGA   819
           || |||||||||||||||||||||||||||||||||||
Db   1680 TAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGGCGTCTGCCACCTTTCTCTGCAGA 1739

Qy   820 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGGAGG   879
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1740 ACTTCGTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGGAGG 1799

Qy   880 CGACTTTCCAGCAGTTCAGAGAGCAGAACGATGTACATAGGCCTTCAAGAGGGAATTGAA   939
           || |||||||||||||||||||||||||||||||||||||||||||||||
Db   1800 CCACTTTCCAGCAGTTCAGAAGCAGAACGACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 1859

Qy   940 AACTAAAGAACCTGTAATCATGAGTACTCTTGACACTGTACGAACTATTTCTGACAGAGCA   999
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1860 AACTAAAGAACCTGTAATCATGAGTACTCTTGACACTGTACGAACTATTTCTGACAGAGCA 1919

Qy   1000 GCCTTTGGAAGGCTAGAGAAACTCTACCAGGAGCACCAGAGAGCTGCCTCCTGAGGGAGAG   1059
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1920 GCCTTTGGAAGGCTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGGAGAG 1979

Qy   1060 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA   1119
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1980 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2039

Qy   1120 AAAATGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGAAAGACT   1179
           |||||||||||||||||||||||||||||||||||||||||||||
Db   2040 AAAATGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGAAAGACT 2099

Qy   1180 CCAGGAACTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGAGGTGAT   1239
           ||||||||||||||||||||||||||||||||||||||||||||
Db   2100 CCAGGAACTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGAGGTGAT 2159

Qy   1240 CAAGGGATCCTGGCAGCCCGTGGGCTACCACTTTGGGCATTCGACTCTTCCAGATCACCTCGA   1299
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2160 CAAGGGATCCTGGCAGCCCGTGGGCTACCACTTTGGGCATTCGACTCTTCCAGATCACCTCGA 2219

Qy   1300 GAAAGTCAAGGACTTCGAGAGAAATTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA   1359
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   2220 GAAAGTCAAGGACTTCGAGAGAAATTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 2279

Qy   1360 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTTGGGCATTCAGCTCTCACCAGATAAACCTCAGCAC   1419
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   2280 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTTGGGCATTCAGCTCTCACCAGATAAACCTCAGCAC 2339

Qy   1420 TCTTGGAAGACCTGAACACCAGACTTCGAGAAGCTTCGCAGGTGAGGACGAGTCAG   1479
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   2340 TCTTGGAAGACCTGAACACCAGACTTCGAGAAGCTTCGCAGGTGAGGACGAGTCAG 2399

Qy   1480 GCAGCTGCATGAGAGCCCCACAGGGACTTTGGTCCAGCATCTTCAGCACTTCTTTCCACGTC   1539
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   2400 GCAGCTGCATGAGAGCCCCACAGGGACTTTGGTCCAGCATCTTCAGCACTTCTTTCCACGTC 2459

Qy   1540 TGTCCAGGGTCCCTGGGAGGAGAGCCATTCTCGCCAAACAAAGTGCCCAACTATCAACCA   1599
           ||||||||||||||||||||||||||||||||||||||||||||
Db   2460 TGTCCAGGGTCCCTGGGAGGAGAGCCATTCTCGCCAAACAAAGTGCCCAACTATCAACCA 2519

Qy   1600 CGAGACTGCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGACTCTACCAGTCTTTAGC   1659
           ||| ||||||||||||||||||||||||||||||||||||
Db   2520 CGAGACTGCAAAACAACTTGCTGGGACCATCCCAAAATGACAGAGACTCTACCAGTCTTTAGC 2579

Qy   1660 TGACCTGAATAATGTCAGATTCTCAGCTTATGTGAATAATGTCAGATTCTCAGCTTATGGACTGCA   1719
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   2580 TGACCTGAATAATGTCAGATTCTCAGCTTATGGACTGCCATGAGAAACTCCGAAGACTGCA 2639

```
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 U; 0 Other;

Query Match             83.3%;  Score 1750.8;  DB 6;  Length 4498;
Best Local Similarity   90.2%;  Pred. No. 0;
Matches 1914;  Conservative  0;  Mismatches 187;  Indels 21;  Gaps 3;
```

```
Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCACCGGAGCCCATT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCACCGGAGCCCATT 1716

Qy    61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAG  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGGAG 1776

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGAGTATTATCGTGCTTCTTTC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGAGTATTATCGTGCTTCTTTC 1836

Qy   181 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1837 TGCTGAGGACACATTGCAAGCACCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy   241 CCAGTTTCATACTCCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGGTTGG  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1897 CCAGTTTCATACTCCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGGTTGG 1956

Qy   301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1957 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA 2016

Qy   361 AACTGAAGTACAAGACAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC  420
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  2017 AACTGAAGTACAAGAGCAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 2076

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTT------TTAATGGATCTCCAGAATCA  474
         |||||||||||||||||||||||||||||||| ||      |||||| |||||||||||||
Db  2077 TAGCATGGAAAAACAAAGCAATTTACATAGAACTCATAGAACTACTGCAACAGTTCCCCCT 2136

Qy   475 GAAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGAAAAATGGA  534
         ||||||||||| |||| |||||||||||| ||||||||||||||||||||||||||||||
Db  2137 GGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGTGAAACAACTGCCAATGTCCTACA 2196

Qy   535 GGAAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAGCGCCAAGTACAACAACATAAGGT  594
         |||| |||| |||  |||||| ||||||||||  ||||||| |||||||||||||||||||
Db  2197 GGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAGAAGAGACTGTAAAAGACGTGATGAA 2256

Qy   595 GCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGT  654
         ||||||||||||  ||| |||||||||| ||||||| ||||| |||||||||||||| |
Db  2257 ACAATGGCAAGAACCTCCAGATGAAATTGAAGCTGAAGGTTTATCACAACCTGGA 2316

Qy   655 GGTAG--------------TGAATGAATCTAGTGGAGAATCACGCAACTGCTGCTTTGGAAGA  702
         |||||              ||||||  ||||||||| ||||||  |||||||||||||||||
Db  2317 TGAAAACAGCCAAAAAATCCTGAGATCCCTGAGAAGGTTCCGATGATGCGAGTCCTGTTACA 2376

Qy   703 ACAACTTAAGG---TATTGGGAGATCGATGGGCAAACATCTGTAGAATGGACAGAAGACCG  759
         |||| |||||    ||||||||||| |||| |||||||||||||  |||||||||||||||
Db  2377 AAGACGTTTGGATAAACATGAACTTCAAGTGCAGTGAGTGAACTTCGGAAAAAGTCTCAACAT 2436

Qy   760 CTGGGGTTCTTTTTACAAGACAGTTCTGACCAGTGGAAGCGTCTGCCACCTTTCCTGCAGGA  819
         |||| ||||  || |||||| ||||| ||||||||||||| |||||||||||||||||||||
Db  2437 TAGGTCCATTTGGAAGCCCAGTTCGACCAGTGGAAGCGTCTGCCACCTTTCCTGCAGGA 2496

Qy   820 ACTTCTGGTGTGGCTACAGCTGCAAAGATGAATGAATTAAGCCGGCACCATTGGAGG  879
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2497 ACTTCTGGTGTGGCTACAGCTGCAAAGATGAATGAATTAAGCCGGCACCATTGGAGG 2556
```

```
Qy  1720 GAAGGCCCTTGTTGGATCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA 1779
         |||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  2640 GAAGGCCCTTGCTTGGATCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA 2699

Qy  1780 CAACCTCAAGCAAAATGACCAGCCATGGATATCCTGCAGATTATAATTGTTTGACCAC 1839
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2700 CAACCTCAAGCAAAATGACCAGCCATGGATATCCTGCAGATTATAATTGTTTGACCAC 2759

Qy  1840 TATTTATGACCGGCTGGAGCCAGCAGCAACAATTTGGTCAACGTCCCTCTCCGGCGTGGA 1899
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2760 TATTTATGACCGGCTGGAGCCAGCAGCAACAATTTGGTCAACGTCCCTCTCCGGCGTGGA 2819

Qy  1900 TATGTGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 1959
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2820 TATGTGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGT 2879

Qy  1960 CCTGTCTTTTAAAACTGGCATCATTTCCCGTGTGTAAAGCACATTGGAAGACAAGTACAG 2019
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2880 CCTGTCTTTTAAAACTGGCATCATTTCCCGTGTGTAAAGCACATTGGAAGACAAGTACAG 2939

Qy  2020 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGGCTGGGCCT 2079
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2940 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGGCTGGGCCT 2999

Qy  2080 CCTTCTGCATGATTCTATCCAA 2101
         ||||||||||||||||||||||
Db  3000 CCTTCTGCATGATTCTATCCAA 3021
```

RESULT 5
AD37258
AAD37258 standard; DNA; 4498 BP.
AAD37258;
21-AUG-2002 (first entry)
Adeno-associated virus vector plasmid, AAV-MCK-3531.
Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin; adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD; Becker muscular dystrophy; ds.
Homo sapiens.
Unidentified.
Chimeric.
WO200183595-A2.
08-NOV-2001.
27-APR-2001; 2001WO-US013677.
28-APR-2000; 2000US-0200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.
Example 1; Page 62-63; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domain, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV)

Qy   1960 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGAAGACAAGTACAG 2019
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3637 CCGTCTTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG 3696

Qy   2020 ATACCTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTGACCAGCCGCAGGGCTGGGCCT 2079
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3697 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGACCAGCCGCAGGGCTGGGCCT 3756

Qy   2080 CCTTCTGCATGATTCTATCCAA 2101
            ||||||||||||||||||||||
Db   3757 CCTTCTGCATGATTCTATCCAA 3778

RESULT 6
AAD37237
ID   AAD37237 standard; DNA; 3858 BP.
XX
AC   AAD37237;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 48-49; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;

Query Match   83.0%;   Score 1743;   DB 6;   Length 3858;
Best Local Similarity   85.8%;   Pred. No. 0;
Matches 2101;   Conservative 0;   Mismatches 0;   Indels 348;   Gaps 1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCATT 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCCTACACGGAGCCCATT 959

Qy   880 CGACTTTCCAGCAGTTCAGAAGCAGCAGAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAA 939
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2557 CGACTTTCCAGCAGTTCAGAAGCAGCAGAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAA 2616

Qy   940 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 999
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2617 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2676

Qy   1000 GCCTTGGAAGGACTAGAGAAACTTACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGAGAGAG 1059
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2677 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGAGAGAG 2736

Qy   1060 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGA 1119
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2737 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGA 2796

Qy   1120 AAAATTGAACCTGCACTCCCGCTGACTGGGCAGGAGAAAAATAGATGAGACCCTTGAAAGACT 1179
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2797 AAAATTGAACCTGCACTCCCGCTGACTGGGCAGGAGAAAAATAGATGAGACCCTTGAAAGACT 2856

Qy   1180 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT 1239
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2857 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT 2916

Qy   1240 CAAGGGATCTGGCAGCCCGTGGCGACTCCTCATTGACTTCTCTCCAAGATCACCTCGA 1299
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2917 CAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2976

Qy   1300 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCCAGTCAA 1359
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2977 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAAGGAACGTGAGCCACGTCAA 3036

Qy   1360 TGACCTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC 1419
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3037 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC 3096

Qy   1420 TCTGGAAGACTGAACACCAGAGCGTTCTGCAGGTGGCCGTCCAGGACCGAGTCAG 1479
      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3097 TCTGGAAGACTGAACACCAGAGCGTTCTGCAGGTGGCCGTCCAGGACCGAGTCAG 3156

Qy   1480 GCAGCTTGCATGAGAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 1539
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3157 GCAGCTTGCATGAGAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3216

Qy   1540 TGTCCAGGGTCCCTGGGAGAGAGCAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 1599
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3217 TGTCCAGGGTCCCTGGGAGAGAGCAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3276

Qy   1600 CGAGACTCAAACAACTGCTGGGACCATCCCAAAATGACAGAGCTCTCACCAGTCTTTAGC 1659
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3277 CGAGACTCAAACAACTGCTGGGACCATCCCAAATGACAGAGCTCTCACCAGTCTTTAGC 3336

Qy   1660 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 1719
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3337 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3396

Qy   1720 GAAGGCCCTTTGCTTGGATCCTTGAGCCTGTCAGCTGCTGCATGTGATGCCTTGGACCAGCA 1779
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3397 GAAGGCCCTTTGCTTGGATCCTTGAGCCTGTCAGCTGCTGCATGTGATGCCTTGGACCAGCA 3456

Qy   1780 CAACCTCAAGCAGCAAAATGACCAGCCCATGGATAATCCTGCAGATTATTAATTGTTTGACCAC 1839
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3457 CAACCTCAAGCAGCAAAATGACCAGCCCATGGATAATCCTGCAGATATCCTGCAGATTATTAATTGTTTGACCAC 3516

Qy   1840 TATTTATGACCGCCTGGAGCAAGCACAACAATTTGGTCAACGTCCCTCCGCGTGGA 1899
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3517 TATTTATGACCGCCTGGAGCAAGCACAACAATTTGGTCAACGTCCCTCTCTGCGGTGGA 3576

Qy   1900 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGGAACAGGAGGATCCGTGT 1959
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3577 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGGAACAGGAGGATCCGTGT 3636

```
Qy   793  GAAGCGTCTGCACCTTCCTGAGGACTTCTGGTGTGGCCACAGCTGAAAGATGATGA  852
          |||||   ||||||||| |||||| |||||||||||| ||||||||||||||||||
Db  2040  GAAGCGTCTGCACCTTCCTGCGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA  2099

Qy   853  ATTAAGCCGGCAGGCACCTATTGGAGGGCACTTCCAGCAGTTCAGAGAGCAGAACGATGT  912
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2100  ATTAAGCCGGCAGGCACCTATTGGAGGGCACTTCCAGCAGTTCAGAGAGCAGAACGATGT  2159

Qy   913  ACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA  972
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2160  ACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA  2219

Qy   973  GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAACGAGACTAGAGAAACTTACCACGGA  1032
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2220  GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAACGAGACTAGAGAAACTTACCACGGA  2279

Qy  1033  GCCCAGAGAGTGCCTCCTGAGGAGAGCCCAGAAATTATCACGCGGCTTCTACGAAAGCA  1092
          |||||||||||||||||||||||||||| |||||  ||| |||||||| |||||||
Db  2280  GCCCAGAGAGTGCCTCCTGAGGAGAGCCCCAGAAATGTCACTCGGCTTCTCACGAAAGCA  2339

Qy  1093  GGCTGAGGAGGTCAATACTGAGTGGGAAAAATGAACCTGCACTCCCGACTGGCGAGAG  1152
          |||||||||||||||||||||||||||||||||||| |||||||||||  ||||||||
Db  2340  GGCTGAGGAGGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTCCCGACTGGCGAGAG  2399

Qy  1153  AAAAATAGATGAGAACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTTGA  1212
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  2400  AAAAATAGATGAGAACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTTGA  2459

Qy  1213  CCTCAAGCTCGCGCCAAGCTGAGTGAGTGATCAAGGATCCTGGCACTCCCGACTCCT  1272
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2460  CCTCAAGCTCGCGCCAAGCTGAGTGAGTGATCAAGGATCCTGGCACTCCCGACTCCT  2519

Qy  1273  CATTGACTCTCTCCAAGATCACCTGAGAAAGTCAAGGCACTTCGAGGAGAATTGCGCC  1332
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2520  CATTGACTCTCTCCAAGATCACCTGAGAAAGTCAAGGCACTTCGAGGAGAATTGCGCC  2579

Qy  1333  TCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTCGCGCAGTTACCACTTTGGGCCAT  1392
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2580  TCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTCGCGCAGTTACCACTTTGGGCAT  2639

Qy  1393  TCAGCTCTCACCGTATAACCTCAGCACTCTGAAGAACCTGACACCAGATGAAGCTTCT  1452
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2640  TCAGCTCTCACCGTATAACCTCAGCACTCTGAAGAACCTGACACCAGATGAAGCTTCT  2699

Qy  1453  GCAGGTGGCCGTCCAGAGACCGACGAGTCAGCGAGTGCATGAAGGCCCACAGGGACTTGGTCC  1512
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2700  GCAGGTGGCCGTCGAGGACCAGCGAGTGCATGAAGGCCCACAGGGACTTGGTCC  2759

Qy  1513  AGCATCTCAGCACTTCTTTCCACGTCTCGCAGGGTCCTCGGGGAGAGCCATCTCGCC  1572
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2760  AGCATCTCAGCACTTCTTTCCACGTCTCGCAGGGTCCTCGGGGAGAGCCATCTCGCC  2819

Qy  1573  AAACAAAGTGCCTCACTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA  1632
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2820  AAACAAAGTGCCTCACTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA  2879

Qy  1633  AATGACAGAGCTCTTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG  1692
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2880  AATGACAGAGCTCTTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG  2939

Qy  1693  GACTGCCATGAAACTCCGAAGACTCGAGAAGACTCGCAGGAGCCCTTTGGATCTCTTGAGCC  1752
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2940  GACTGCCATGAAACTCCGAAGACTCGAGAAGACTCGCAGGAGCCCTTTGGATCTCTTGAGCC  2999

Qy  1753  AGCTGCATGTGATGCTTGGCCTGGATGCCTTGGACCAGCACAACAACTTCAAGCCCATGGATAT  1812
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3000  AGCTGCATGTGATGCTTGGCCTGGATGCCTTGGACCAGCACAACAACTTCAAGCCCATGGATAT  3059

Qy  1813  CCTGCGAGATTATTAATTGTTGACCACTATTTATGTTGACCCTGAGCAAGAGCACAACAA  1872
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3060  CCTGCAGATTATTAATTGTTGACCACTATTTATGTTGACCCTGAGCAAGAGCACAACAA  3119

Qy  1873  TTTGGTCAACGTCCCTCTCTCGCCTGGATATGTGTCTGAACTGCTGCTGAATGTTTATGA  1932
```

```
Qy    61  TCCTTCACAGCATTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGGATGATGGAGAG  120
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   960  TCCTTCACAGCATTGAAGCTCCTGAAGACAACTCATTTGGCAGTTCATTGGATGATGGAGAG  1019

Qy   121  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC  1079

Qy   181  TGCTGAGGACACATTGCAAGCAGGAGAGATTCTAATGATGTGAAGTGGTGGTGAAAGA  240
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  1080  TGCTGAGGACACATTGCAAGCACGAGGAGATTCTAATGATGTGAAGTGGTGGTGAAAGA  1139

Qy   241  CCAGTTCATATCTCATGAGGGGTACATGATGATTGACAGCGCCATCAGGGCCGGGGTTGG  300
          |||||||||||||||||||||||||||||||||||||||| ||||||| |||||||||||
Db  1140  CCAGTTCATATCTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGGCCGGGGTTGG  1199

Qy   301  TAATATTCCACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATTATCAGAAGATGAAGA  360
          |||||||| ||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  1200  TAATATTCATACAATTGGGAAGTAAGCTGATTGGAAACACGAAAATTATCAGAAGATGAAGA  1259

Qy   361  AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGAATGGGAATGCCTCAGGGTAGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260  AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAGAATGGGAATGCCTCAGGGTAGC  1319

Qy   421  TAGCATGGAAAACAAAGCAATTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  480
          ||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1320  TAGCACATGGAAAAACAAAGCAATTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  1379

Qy   481  GAAAGAGTTGAATGCTAACAGAAACAGAGAGAAAGACAAGGAAAAATGGAGGAAGA  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1380  GAAAGAGTTGAATGCTAACAGAAACAGAGAGAAAGACAAGGAAAAATGGAGGAAGA  1439

Qy   541  GCCTCTTGAAGACCTGATCTTGAAGACTACAAAACGGCCAAGTACAACAACATAAGGTGCTTCA  600
          |||||||  ||||||||||||||||||||  |||||||||||||||||||||||||||||
Db  1440  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACAACATAAGGTGCTTCA  1499

Qy   601  AGAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACAACTTAGTAGT  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1500  AGAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACAACTTAGTAGT  1559

Qy   661  TGATGAATCTCAGTGGTGGAGATCACGGCAACAACTGCTCTCTGCTGCTTGGAAGAACTTAAGGTATTGGG  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1560  TGATGAATCTCAGTGGTGGAGATCACGGCAACAACTGCTCTCTGCTGCTTGGAAGAACTTAAGGTATTGGG  1619

Qy   721  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGA--  777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1620  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC  1679

Qy   778  --------------------------------------------------------  777
          
Db  1680  TCATAGATTACTGCAACAGTCCCCTGGACCTGAAAAGTTTCTTGCTGGCTTACAGA  1739

Qy   778  --------------------------------------------------------  777
          
Db  1740  AGCTGAAACAACTGCCCAATGTCCTACAGATGCTACCCGTAAGGAAAGGCTCCTGAGAAGA  1799

Qy   778  --------------------------------------------------------  777
          
Db  1800  CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTTCCAAGGTGAAATTGAAGC  1859

Qy   778  --------------------------------------------------------  777
          
Db  1860  TCACCAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAAGATCCCTGGA  1919

Qy  1920  AGGTTCCGATGATGATGCCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTTGGAG  1979
          
Db   777  --------------------------------------------------------

Qy   778  ------------CAGTTCTGACCAGTG  792
          
Db  3060  --------------------------------------------------------

Qy  1980  TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCGACCCAGTG  2039
```

```
Qy      1 GAGCTATGCTACACACCAGGCTGCTTATGTCACCACCTTCTGACCTCTGACGAGGAGCCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCTACACACCAGGCTGCTTATGTCACCACCTTCTGACCTCTGACGAGGAGCCATT 1716

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1776

Qy    121 TGAAGTAAACCTTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy    241 CCAGTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGCCGGGTTGG 1956

Qy    301 TAATATTCTACAATTGGGAAGCTGATTGGAACAGCAGGAATGAAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGCTGATTGGAACAGCAGGAATGAAAATTATCAGAAGATGAAGA 2016

Qy    361 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

Qy    421 TAGCATGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACT 2136

Qy    481 GAAAGAGTGAATGACTGGCTAACAAAAACAGAAGAAACAGAAGAAATGGAGGAAGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTGAATGACTGGCTAACAAAAACAGAAGAAACAGAAGAAATGGAGGAAGA 2196

Qy    541 GCCTCTTGGACCTGACCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCTCTTGGACCTGACCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 2256

Qy    601 AGAAGATCTAGAACAAGAAGTCAGGGTCAATTCTCCACTCCATGGTGGTGGTAGT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 AGAAGATCTAGAACAAGAAGTCAGGGTCAATTCTCCACTCCATGGTGGTGGTAGT 2316

Qy    661 TGATGAATCTAGTGGAGATCACGCAACTGCTTTGGAAGAACAACTTAAGGTATTGGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TGATGAATCTAGTGGAGATCACGCAACTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376

Qy    721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGA-- 777
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 2436

Qy    778 ------------------------------------------------------------- 777
Db   2437 TCATAGATTACTGCAACAGTTCCACCAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGA 2496

Qy    778 ---------------------------------------------------------- 777
Db   2497 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCGTCCTAGAAGA 2556

Qy    778 ------------------------------------------------------------- 777
Db   2557 CTCCAAGGGAGTAAAAGAGTTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2616

Qy    778 ---------------------------------------------------------- 777
Db   2617 TCACACAGATGTTTATCACACCTGATGAAAACAGCCAAAAATCCTGAGATCCCTGGA 2676

Qy    778 ----------------------------------------------------------- 777
Db   2677 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGGAG 2736
```

```
Db   3120 TTTGGTCAACGTCCCTCTCGCCTGGCGTGGATATGTGCTGAACTGGCTGCTGAATGTTTATGA 3179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1933 TACGGGACGAACAGGAGGGAGGATCCGTCCTGTCCTTCTTTAAAACTGGCATCATTTCCCTGTG 1992

Db   3180 TACGGGACGAACAGGAGGGAGGATCCGTCCTGTCCTGTCCTTCTTTAAAACTGGCATCATTTCCCTGTG 3239
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1993 TAAAGCACATTTGGAAGCAAGCAAGTACAGATACCTTTTCAAGCAAGTGCCAAGTTCAACAGG 2052

Db   3240 TAAAGCACATTTGGAAGCAAGCAAGTACAGATACCTTTTCAAGCAAGTGCCAAGTTCAACAGG 3299
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2053 ATTTTGTGACCAGCGCAGGCTGGCCTCCCTTCTGCATGATTCTATCCAA 2101

Db   3300 ATTTTGTGACCAGCGCAGGCTGGCCTCCCTTCTGCATGATTCTATCCAA 3348
```

RESULT 7
AD37257
ID AD37257 standard; DNA; 4825 BP.
XX
AC AAD37257;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013677.
XX
PR 28-APR-2000; 2000US-0200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
PR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT gene.
XX
PS Example 1; Page 61-62; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;

Query Match   83.0%; Score 1743; DB 6; Length 4825;
Best Local Similarity   85.8%; Pred. NO. 0;
Matches 2101; Conservative   0; Mismatches   0; Indels 348; Gaps   1;

```
RESULT 8
AAD37263
ID   AAD37263  standard; DNA; 4848 BP.
XX
AC   AAD37263;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Cytomegalovirus.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 68-70; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a
CC   cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
```

SQ   Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;

Query Match   83.0%;   Score 1743;   DB 6;   Length 4848;
Best Local Similarity 85.8%;   Pred. No. 0;   Indels 348;   Gaps 1;
Matches 2101;   Conservative 0;   Mismatches 0;

```
Qy        1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTCTGACCCTACACGGAGCCCATT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1680 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTCTGACCCTACACGGAGCCCATT 1739

Qy       61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTCGGCAGTTCATTGATGAGAGAG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1740 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTCGGCAGTTCATTGATGAGAGAG 1799

Qy      121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC    180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1800 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   1859

Qy      181 TGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA    240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1860 TGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   1919

Qy      241 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG    300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1920 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTGG   1979

Qy      301 TAATATTCACAATTGGGAAGCTGATTGGAACAGGAGAAAATTATCAGAGATGAAGA      360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1980 TAATATTCACAATTGGGAAGCTGATTGGAACAGGAGAAAATTATCAGAGATGAAGA     2039

Qy      361 AACTGAAGTACAAGACAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2040 AACTGAAGTACAAGACAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 2099

Qy      421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2100 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  2159

Qy      481 GAAAGAGTTGAATGCTACTACAAAAACAGAGAAGAAGGAAAAATGGAGGAAGA        540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2160 GAAAGAGTTGAATGCTACTACAAAAACAGAGAAGAAGGAAAAATGGAGGAAGA       2219

Qy      541 GCCTCTTGGACCTGATCTTGAAGACTTTGAAGACCTAAAACGCCAGTACACAACATAGGTGCTTCA  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2220 GCCTCTTGGACCTGATCTTGAAGACTTTGAAGACCTAAAACGCCAGTACACAACATAGGTGCTTCA 2279

Qy      601 AGAAGATCTAGAACAAGAACAAGACAAGTCAGGGTCAATTCTCTACTCACATGGTGGTAGT  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2280 AGAAGATCTAGAACAAGAACAAGACAAGTCAGGGTCAATTCTCTACTCACATGGTGGTAGT 2339

Qy      661 TGAATGAATCTAGTGGAGATCACCGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2340 TGAATGAATCTAGTGGAGATCACCGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 2399

Qy      721 AGATCGATGGCAAACATCTGTAGATGAGACAGAAGAACCGCTGGGTTCTTTTTTACAAGA--  777
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2400 AGATCGATGGCAAACATCTGTAGATGAGACAGAAGAACCGCTGGGTTCTTTTTTACAAGACAC 2459

Qy      778 ---------------------------------------------------------------  777
Db     2460 TCATAGATTACTGCACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTTACAGA         2519

Qy      778 ---------------------------------------------------------------  777
Db     2520 AGCTGAAACAACTGCCCTACTACAGTCTTACCACCCGTAAGGAAAGGCTCCTAGGAAGA     2579

Qy      778 ---------------------------------------------------------------  777
Db     2580 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCAAGACCCTCCAAGGTGAAATTGAAGC     2639

Qy      778 ---------------------------------------------------------------  777
Db     2640 TCACACAGATGTGTTTATCAACAACTGGATGAAAAACAGCCAAAAAAATCCTGAGATCCCTGGA 2699

Qy      778 ---------------------------------------------------------------  777
Db     2700 AGGTTCCGATGATGCAGTCCTGTTACCAAACGCGTTTGGGATAACATGAGAACTTCAGTGGAG 2759

Qy      778 ---------------------------------CAGTTCTGACCAGTG  792
                                             ||||||||||||||
Db     2760 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCATTTGGAAGCCAGTTCTGACCAGTG 2819

Qy      793 GAAGCGTCTGCACCTTTCTTCTGCGACTTCTGGTGTGGCTACAGCTGAAAGATGATGA  852
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2820 GAAGCGTCTGCACCTTTCTTCTGCGACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2879

Qy      853 ATTAAGCCGGCACCTATTGGAGGCGACTTTCAGAAGACAGAACGACATGT  912
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2880 ATTAAGCCGGCACCTATTGGAGGCGACTTTCAGAAGACAGAACGACATGT 2939

Qy      913 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAAACCTGTAATCATGAGTACTCTTGA  972
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2940 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAAACCTGTAATCATGAGTACTCTTGA 2999

Qy      973 GACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1032
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3000 GACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 3059

Qy     1033 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGATGTCACTCGGCTTCTACGAAAGCA 1092
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3060 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGATGTCACTCGGCTTCTACGAAAGCA 3119

Qy     1093 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGACTCCGCTACCTCGGCCAGAG 1152
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3120 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGACTCCGCTACCTCGGCCAGAG 3179

Qy     1153 AAAAATAGATGAGACCCTGAAAGACTCCAGAGACTCAAGAGGCCACCGGATGAGCTGGA 1212
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3180 AAAAATAGATGAGACCCTGAAAGACTCCAGAGACTCAAGAGGCCACCGGATGAGCTGGA 3239

Qy     1213 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCGGCAGCCCGTGGGCGGATCTCCT 1272
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3240 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCGGCAGCCCGTGGGCGGATCTCCT 3299

Qy     1273 CATTGACTTCTCTCTCCAAGATCAAGATGACCTTGCTCGCCCAGCTTACCACTTGGGCAT 1332
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3300 CATTGACTTCTCTCTCCAAGATCAAGATGACCTTGCTCGCCCAGCTTACCACTTGGGCAT 3359

Qy     1333 TCTGAAAGAGAACGTGAGCGACACGTCACGTCAGGCAGTGCAGCGTCAGGCAGGACTTCT 1392
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3360 TCTGAAAGAGAACGTGAGCGACACGTCACGTCAGGCAGTGCAGCGTCAGGCAGGACTTCT 3419

Qy     1393 GCAGGTGGCCGTGGAGGACCGAGTCAGGCAGCGAGCAGCCCACAGGGAGAGCCATCCGCC 1452
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3420 GCAGGTGGCCGTGGAGGACCGAGTCAGGCAGCGAGCAGCCCACAGGGAGAGCCATCCGCC 3479

Qy     1453 AGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGGTCCCTGGGACTTGGACCATCCGCC 1512
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3480 AGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGGTCCCTGGGACTTGGACCATCCGCC 3539

Qy     1513 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA 1572
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3540 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA 3599

Qy     1573 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 1632
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3600 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3659

Qy     1633 GACTGCCATGAAACTCCGAAGACTGCAGAAGCTGCAGAAGTGAAATTCTCAGCTTATAG 1692
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3660 GACTGCCATGAAACTCCGAAGACTGCAGAAGCTGCAGAAGTGAAATTCTCAGCTTATAG 3719

Qy     1693 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGGAGCTTGTC 1752
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3720 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGGAGCTTGTC 3779
```

```
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a
XX   small polyA signal sequence
SQ   Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

     Query Match             83.0%;  Score 1743;  DB 6;  Length 5060;
     Best Local Similarity   85.8%;  Pred. No.0;
     Matches 2101;  Conservative   0;  Mismatches   0;  Indels 348;  Gaps    1;

Qy       1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1892 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1951

Qy      61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1952 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 2011

Qy     121 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAGAAGAGTATTATCGTGGCTTCTTTC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2012 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAGAAGAGTATTATCGTGGCTTCTTTC 2071

Qy     181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2072 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 2131

Qy     241 CCAGTTCATACTCATGAGGGGTACATGATGGGATTTGACAGCCCATCAGGGCCGGGGTTGG  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2132 CCAGTTCATACTCATGAGGGGTACATGATGGGATTTGACAGCCCATCAGGGCCGGGGTTGG 2191

Qy     301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAGATGAAGA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2192 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAGATGAAGA 2251

Qy     361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2252 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2311

Qy     421 TAGCATGGAAAAACAAAGCAATTACATTACATAGAGTTTAATGAGTTTAATGATCTCCAGAATCAGAAACT  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2312 TAGCATGGAAAAACAAAGCAATTACATTACATAGAGTTTAATGAGTTTAATGATCTCCAGAATCAGAAACT 2371

Qy     481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGAGAGAAGA  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAAAATGAGAGAAGA 2431

Qy     541 GCCTCTTGGACCTGGATCTTGAAGACCTAAAACGGCCAAGTACAACACATAAGGTGCTTCA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2432 GCCTCTTGGACCTGGATCTTGAAGACCTAAAACGGCCAAGTACAACACATAAGGTGCTTCA 2491

Qy     601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTTCACTCACATGGTGTGGTAGT  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2492 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTTCACTCACATGGTGTGGTAGT 2551

Qy     661 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTGGAAGAACAACTAAGGTATTGGG  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2552 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTGGAAGAACAACTAAGGTATTGGG 2611

Qy     721 AGATCATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTACAGA---  777
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2612 AGATCATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTACAAGACAC 2671

Qy     778 ------------  777
Db    2672 TCATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCTGGCTTACAGA 2731

Qy     778 ------------  777
Db    2732 AGCTGAAAACAACTGCCAATGTCCTACAGATGTCTACCCGTAAGGAAAGGCTCCTAGAGA 2791

Qy     778 ------------  777
```

```
     1753 AGCTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATAT 1812
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     3780 AGCTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATAT 3839

     1813 CCTGCAGATTATTAATTGTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAACAA 1872
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     3840 CCTGCAGATTATTAATTGTTGACCACTATTATGACCGCCTGGAGCAAGAGCACAACAA 3899

     1873 TTTGGTCAACGTCCCTCCTCTCCGCGGTGGTGATATGTGTGAACTGGCTGCTGAATGTTTATGA 1932
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     3900 TTTGGTCAACGTCCCTCCTCTCCGCGGTGGTGATATGTGTGAACTGGCTGCTGAATGTTTATGA 3959

     1933 TACGGGACGAACAGGAGGAGGATCCCGTGTCCGTGTCTTTTAAAACTGGCATCATTTCCCTGTG 1992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     3960 TACGGGACGAACAGGAGGAGGATCCCGTGTCCGTGTCTTTTAAAACTGGCATCATTTCCCTGTG 4019

     1993 TAAAGCACATTGGAAGACAAGAACACCTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
     4020 TAAAGCACATTGGAAGACAAGAACACCTTTCAAGCAAGTGGCAAGTTCAACAGG 4079

     2053 ATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCGCATGCATGATTCTATCCAA 2101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
     4080 ATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCGCATGCATGATTCTATCCAA 4128
```

```
ESULT 9
AD37264

DD   AAD37264 standard; DNA; 5060 BP.
CC   AAD37264;
CC
TT   21-AUG-2002  (first entry)
BB   Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
XX
XX   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX   Becker muscular dystrophy; ds.
SS
SS   Homo sapiens.
SS   Cytomegalovirus.
SS   Unidentified.
SS   Chimeric.
NN
NN   WO200183695-A2.
DD   08-NOV-2001.
XX
FF   27-APR-2001; 2001WO-US013677.
RR   28-APR-2000; 2000US-0200777P.
AA
AA   (XIAO/) XIAO X.
XX   Xiao X;
II
RR   WPI; 2002-049342/06.
RR
TT   New dystrophin minigene for treating Duchenne or Becker muscular
TT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
TT   rod repeats, utrophin or a spectrin gene, H1 and H4 domains and a cysteine rich domain of a dystrophin
TT   gene.
XX
XX   Example 1; Page 70-71; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
```

```
Db  3872 AATGACAGAGCTACCAGTGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3931
Qy  1693 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTGCTTGGATCTTCTTGAGCCTGTC 1752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3932 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTGCTTGGATCTTCTTGAGCCTGTC 3991
Qy  1753 AGCTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATAT 1812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3992 AGCTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATAT 4051
Qy  1813 CCTGCAGATTATTAATTGTTGACCACTATTATGACCGGCCTGGAGCAAGAGCACAACAA 1872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4052 CCTGCAGATTATTAATTGTTGACCACTATTATGACCGGCCTGGAGCAAGAGCACAACAA 4111
Qy  1873 TTTGGTCAACGTCCCTCCTCCGCGTGGATATGTGTCTGAACTGGCTGGCTGTGAATGTTTATGA 1932
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4112 TTTGGTCAACGTCCCTCCTCCGCGTGGATATGTGTCTGAACTGGCTGGCTGTGAATGTTTATGA 4171
Qy  1933 TACGGGACGAACAGGGAGGATCCGTGTCTCGTGTCCTGTCTCTTTAAAACTGGCATCATTTCCCTGTG 1992
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4172 TACGGGACGAACAGGGAGGATCCGTGTCTCGTGTCCTGTCTCTTTAAAACTGGCATCATTTCCCTGTG 4231
Qy  1993 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4232 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGG 4291
Qy  2053 ATTTTGTGACCAGCCAGGCTGGGCCTCCTTCTGCATGATTCTATCCAA 2101
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db  4292 ATTTTGTGACCAGCCAGGCTGGGCCTCCTTCTGCATGATTCTATCCAA 4340
```

```
RESULT 10
AAD37260
ID   AAD37260 standard; DNA; 4414 BP.
XX
AC   AAD37260;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-3447.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 65-66; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
```

```
.b  2792 CTCCAAGGGAGTAAAAGAGCTGATGAAACATGGCAAGACCTCCAAGGTGAAATTGAAGC 2851
y    778 ------------------------------------------------------------ 777
.b  2852 TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 2911
y    778 ------------------------------------------------------------ 777
.b  2912 AGGTTCCGATGATGCAGTCAGTCCTGTTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 2971
y    778 ---------------------------CAGTTCTGACCAGTG 792
                                     |||||||||||||||
.b  2972 TGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTG 3031
y    793 GAAGCGGCTGACCACCTTTCTCTGCAGAACTTCTGTGTGGCTACAGCTGAAAGATGATGA 852
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3032 GAAGCGGCTGACCACCTTTCTCTGCAGAACTTCTGTGTGGCTACAGCTGAAAGATGATGA 3091
y    853 ATTAAGCCGGCAGGCACCTATTGGAGGCGACCTTCCAGCAGTTCAGAAGCAGAACGATGT 912
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3092 ATTAAGCCGGCAGGCACCTATTGGAGGCGACCTTCCAGCAGTTCAGAAGCAGAACGATGT 3151
y    913 ACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAAACTGTAATCATGAGTACTCCTTGA 972
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3152 ACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAAACTGTAATCATGAGTACTCCTTGA 3211
y    973 GACTGTACGAAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGA 1032
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3212 GACTGTACGAAATATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACTCTACCAGGA 3271
y   1033 GCCCAGAGAGAGTGCCTCCTGAGGAGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1092
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3272 GCCCAGAGAGAGTGCCTCCTGAGGAGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCA 3331
y   1093 GGCTGAGGAGGTCAATACTGATGTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 1152
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3332 GGCTGAGGAGGTCAATACTGATGTGAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAG 3391
y   1153 AAAAATAGAGATGAGACCCTTGAAAAGACTCCAGGAACTCAAGAGAGGCCACGGATGAGCTGGA 1212
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3392 AAAAATAGAGATGAGACCCTTGAAAAGACTCCAGGAACTCAAGAGAGGCCACGGATGAGCTGA 3451
y   1213 CCTCAAGCTGCCCGCCAAGCTGCCGAGGGTGATCAAGGGATCAAGAGGCAGCCCGTGGGCCCGATCCT 1272
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3452 CCTCAAGCTGCCCGCCAAGCTGCCGAGGGTGATCAAGGGATCAAGAGGCAGCCCGTGGGCCCGATCCT 3511
y   1273 CATTGACTCTCTCCAAGATCACCTCCAGATCACCTGGGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1332
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3512 CATTGACTCTCTCCAAGATCACCTCCAGATCACCTGGGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 3571
y   1333 TCTGAAAGAGAACGTGAGCGACGTGAGCCGACGTTGCTCGGCCAGTTACCACTTTGGGCAT 1392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3572 TCTGAAAGAGAACGTGAGCGACGTGAGCCGACGTTGCTCGGCCAGTTACCACTTTGGGCAT 3631
y   1393 TCAGCTCTCACCGTATAACCTCCGCAGCACTCTCGGAGACACCTGAACACCAGATGGAAGCTTCT 1452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3632 TCAGCTCTCACCGTATAACCTCCGCAGCACTCTCGGAGACACCTGAACACCAGATGGAAGCTTCT 3691
y   1453 GCAGGTGGCCGTTCGAGGACCGAGTCAGGCAGCGAGTGCGCATGAGAAGCCCACAGGGACTTTTGGTCC 1512
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3692 GCAGGTGGCCGTTCGAGGACCGAGTCAGGCAGCGAGTGCGCATGAGAAGCCCACAGGGACTTTTGGTCC 3751
y   1513 AGCATCCAGCAGCACTTTCTTTTCCACGTTCCCAGGGTGTCCCTGGGAGAGAGCCATCTCGCC 1572
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b  3752 AGCATCCAGCAGCACTTTCTTTTCCACGTTCCCAGGGTGTCCCTGGGAGAGAGCCATCTCGCC 3811
y   1573 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACATCCCAA 1632
           ||||||||||||||||||||||||||||||||||||||||||||||
.b  3812 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACAACTTGCTGGGACCATCCCAA 3871
y   1633 AATGACAGAGCTCTCACCAGTGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 1692
```

domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is AAV
vector plasmid construct containing human dystrophin minigenes, a muscle
creatine kinase (MCK) promoter and a small polyA signal sequence

Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 U; 0 Other;

Query Match 78.1%; Score 1641; DB 6; Length 4414;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1848; Conservative 0; Mismatches 190; Indels 63; Gaps 3;

```
Qy      1 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT     60
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1657 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   1716

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGAGAG    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGAGAG   1776

Qy    121 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTCTTTC   180
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTCTTTC  1836

Qy    181 TGCTGAGGACACATTGCAAGCAACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA   240
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCAACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  1896

Qy    241 CCAGTTCATATACTCATGAGGGGTACAGTAAGCTGATTGACAGCCATCAGGGCCGGGTTGG   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTCATATACTCATGAGGGGTACAGTAAGCTGATTGACAGCCATCAGGGCCGGGTTGG  1956

Qy    301 TAATATTCACAATTGGGAAGTAAGCTGATTGAACAGGAAAATTATCAGAAGATGAAGA    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCACAATTGGGAAGTAAGCTGATTGAACAGGAAAATTATCAGAAGATGAAGA   2016

Qy    361 AACTGAAGTACAAGAGCAGATGAATCAATTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGATGAATCAATTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC  2076

Qy    421 TAGCATGAGTTGAAGATCAGATGAATCAATTACTATAGAGTTTAATGAGTTTAATGATCTCCAGAAACT   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGAGTTGAAGATCAGATGAATCAATTACTATAGAGTTTAATGAGTTTAATGATCTCCAGAAACT  2136

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA    540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAGAACAAGGAAAATGGAGGAAGA   2196

Qy    541 GCCTCTTGGACCTGATCTTGAAGACAAGCGGGTCAAGGTACAACATAAGGTGCTTCA    600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCTCTTGGACCTGATCTTGAAGACAAGCGGGTCAAGGTACAACATAAGGTGCTTCA   2256

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTGCTGCTTGCTGTGTGGTGTAGT   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTGCTGCTTGCTGTGTGGTGTAGT  2316

Qy    661 TGATGAATCTAGTTGGGACAAACATCTGAGATGGCAGAACTGGTTCAAGAGAACTTAAGGTATTGGG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TGATGAATCTAGTTGGGACAAACATCTGAGATGGCAGAACTGGTTCAAGAGAACTTAAGGTATTGGG  2376

Qy    721 AGATCGATGGGCAAACATCTGAGATGACAGTCGCGGTTCTTTTCTTTTACAGACAG    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 AGATCGATGGGCAAACATCTGAGATGACAGTCGCGGTTCTTTTCTTTTACAGACAG   2436

Qy    781 TTCTGACCAGTGGAAGCGTCTGGACCTTTCTGACCAGTGGTGTGGCTACAGCT    840
          ||||||||||||||||||| |||||||||||||| |||||||||||||||||
Db   2437 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTAGCATGGCTTTTCAGA    2496

Qy    841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAA   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2497 AAAAGAAGATGCAGTGAACAAGAGATTCACACACTGGCTTTAAAGATCAAAAT GAAATGTT  2556

Qy    901 GCAGAACGATGTACATGAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAAACCTGTAATCAT    960
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   2557 ATCAAGTCTTCAAAAATCTGGCCGTTTTAAAAGCGGATCTGAAAGAGAAAAAGCAATCCAT  2616

Qy    961 GAGTACTTGAGACTGTACGAATATTTCTGACAGCAGCCTTTGGAAGGACTAGAGAA   1020
          ||||||||| | |   || ||||  |||   |||||| | |  ||||||
Db   2617 GGGCAAACTGTA----------TTCACTCAAACAAGATCTTCTTTCAACACTGAAG   2662

Qy   1021 ACTCTACCAGAGGAGCCCAGAGAGCTGCCTCCTGCAGGAGAGAGCCCAGAATGTCACTCGGCT   1080
          ||||||||||||||||| || |||| |||||||||||||||||||||||||||
Db   2663 AATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGATAACTTGCCGGGTGTTGG--  2720

Qy   1081 TCTACGAAAGCACAGGCTGAGGAGTCAATACTGATGGAAAAATTGAACCTGCCACTCCCGC   1140
          |||||||||||||||||||||||| | |    || |||
Db   2721 --------------------------ATAATTAGTCCAAAACTGA-----------------  2743

Qy   1141 TGACTGGCAGAGAAAAATAGATCAGACCCTTGAAAGACTTCCAGAACTTCAAGAGGCCAC   1200
          |||||||||||||  |  || || ||||||||||||||||||||||||||||||||
Db   2744 --------AAGATACAGCACAGACCCTTGAAAGACTTCCAGAACTTCAAGAGGCCAC   2793

Qy   1201 GGATGAGCTGGACCTCAAGCTGCGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGT   1260
          |||||||||||||||||||||| || |||||||||||||||||||||||||||||||||
Db   2794 GGATGAGCTGGACCTCAAGCTGCGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGT  2853

Qy   1261 GGGCGATCTCCTCATTGACTTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG   1320
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   2854 GGGCGGATCTCCTCATTGACTTCTCCAAGATCACCTCGAGAAGTCAAGGCACTTCGAGG  2913

Qy   1321 AGAAATTGCGCCTCTGAAGAAGAACGTGAGCCACGTCAATGAACCTTGCTCGCCAGCTTAC   1380
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   2914 AGAAATTGCGCCTCTGAAGAAGAACGTGAGCCGCAGTCAATGACCTTGCTCGCCAGCTTAC  2973

Qy   1381 CACTTTGGGCATTCAGCTCTCAGGACTCTCAGGACCGCTCAGAACTGCTGGAAGATGAAGA   1440
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   2974 CACTTTGGGCATTCAGCTCTCAGGACTCTCAGGACCGCTCAGAACTGCTGGAAGATGAAGA  3033

Qy   1441 ATGGAAGCTTCTCGCAGCATCCAGGACTTCTTCTTCCAGTGTCCGCAGCTGCAGGAGAG   1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3034 ATGGAAGCTTCTCGCAGCATCCAGGACTTCTTCTTCCAGTGTCCGCAGCTGCAGGAGAG  3093

Qy   1501 GGACTTTGTTGGTCCAGGTGGCCATCCAGGACTTCTTTCCAGTGTCGTCCAGGGTCCCTGGGAGAG   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3094 GGACTTTGTTGGTCCAGGTGGCCATCCAGGACTTCTTTCCAGTGTCGTCCAGGGTCCCTGGGAGAG  3153

Qy   1561 AGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACAGACTCAAACAACTGCTG   1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3154 AGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACAGACTCAAACAACTGCTG   3213

Qy   1621 GGACCATCCCAAAATGACAGAGCTCTTAGCTGACCTGGAATAATGTCAGATT   1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3214 GGACCATCCCAAAATGACAGAGCTCTTAGCTGACCTGGAATAATGTCAGATT   3273

Qy   1681 CTCAGCTTATAGAGACTGCCATGAAACTCCGAGACTGCAGAGAGGCCCTTTGCTTGGATCT   1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3274 CTCAGCTTATAGAGACTGCCATGAAACTCCGAGACTGCAGAGAGGCCCTTTGCTTGGATCT  3333

Qy   1741 CTTGAGCCTGCAGTGATGCCATGGCTTGATGTGCCTTGACCAGCCACTATTAATGTTTGAGCA   1800
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   3334 CTTGAGCCTGCAGTGATGCCATGGCTTGATGTGCCTTGACCACCTATTAATGTTTACAAGACAG  3393

Qy   1801 GCCCATGGATATCCTGCAGATTATTAGAGATGGCCCTGGTTCCTTTACTATTTGGG   1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3394 GCCCATGGATATCCTGCAGATTATTAGAGATGGCCCTGGTTCCTTTTCTTTACAGACACAT  3453

Qy   1861 AGAGCAACACAATTGGTCAACGTCTCCTCTCTGCGTGTGCTGGCTACAGCT   1920
          |||||||||| |||||||||||||| |||||||| ||
Db   3454 AGAGCAACACAATTGGTCAACGTCTCCTCTCTGCGTGTGCTGGCTACAGCT   3513

Qy   1921 GAATGTTTATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAA   1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        (continued)
```

```
Db  3514  GAATGTTTATGATACGGGACGAACAGGGAGGAGGATCCGTGCTGCCTGTCTTTAAAACTGGCAT  3573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   900  GAGCTATGCCTACACACCAGGCTGCTTATGTCCACCACCTTCGACCCCTACACGGAGCCCATT   959

Qy  1981  CATTTCCCTGTGTAAAGCACATTTGAAGACAAGACAAGACAGATACCTTTTCAAGCAAGTGGC  2040
          |||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db  3574  CATTTCCCTGTGTAAAGCACATTTGGAAGCTCCTGAAGACAAGCAAGTCATTTGGCAGTTCATTGATGGAGAG  3633
Qy    61  TCCTTCACAGATTTGGAAGCTCCTGAAGACAAGCAAGTCATTTGGCAGTTCATTGATGGAGAG   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960  TCCTTCACAGATTTGGAAGCTCCTGAAGACAAGCAAGTCATTTGGCAGTTCATTGATGGAGAG  1019

Qy  2041  AAGTTCAACAGAGTTTGTGACCAGCGCAGGCTGGGCCTCCTTCGCATGATTCTATCCA  2100
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  3634  AAGTTCAACAGAGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCGCATGATTCTATCCA  3693
Qy   121  TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGAAGTATTATCGTGGCTTCTTTC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020  TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAAGAAGTATTATCGTGGCTTCTTTC  1079

Qy  2101  A  2101
          |
Db  3694  A  3694
Qy   181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1080  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  1139

Qy   241  CCAGTTTCATACTCATGAGGGGTACAGATGATGATTGACAGCCCATCAGGGCCGGGTTGG   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1140  CCAGTTTCATACTCATGAGGGGTACAGATGATGATTGACAGCCCATCAGGGCCGGGTTGG  1199

Qy   301  TAATATTCTACAATTGGGAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200  TAATATTCTACAATTGGGAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA  1259

Qy   361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC  1319

Qy   421  TAGCATGAAAAACAAGACCATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT   480
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  1320  TAGCATGAAAAACAAGACCATTTACATAGAGTTTTAATGGATCTCCAGAATC-GAAACT  1378

Qy   481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGAGAAATGGGAGGAAGA   540
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  1379  GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAGAACAAGAGAAATGGGAGGAAGA  1438

Qy   541  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA   600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1439  GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA  1498

Qy   601  AGAAGATCTAGAACAAGACAAGAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTCGTAGT   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1499  AGAAGATCTAGAACAAGACAAGAGTCAGGGTCAATTCTCTCCACTCACATGGTGGTCGTAGT  1558

Qy   661  TGATGAATCTAGTTGGAGATCACGCAACTGCTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1559  TGATGAATCTAGTTGGAGATCACGCAACTGCTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG  1618

Qy   721  AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAG   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1619  AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAT  1678

Qy   781  TTCTGACCAGTGAAGGGTCTCGCACCTTTCTACTGGCAACAGTGTCCTTTTTAGTTAGTGGCATGCATGCTTTCAGA   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1679  CCTTCTCAAATGGCAAGTGGCAACGTCTACTGAAGAACAGTGTCCTTTTAGTTAGTGGCATGCATGGCTTTCAGA  1738

Qy   841  GAAAGATGATGAATTAAGCCAGCGCCAGGCGACCATTCCAGCAGTTCCAGATTCCAGAA   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1739  AAAGAAGATGCAGTGACAAGATTCACACACTGGCTTTAAAGATCAAAATGAAAATGTT  1798

Qy   901  GCAGAACGAATGTACATAGGGGCCTTCAGAGGGGAATTGAACTAAAGAACCTGTAATCAT   960
          |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db  1799  ATCAAGTCTTCAAAAACTGGCCGGTTTTAAAACGGGAATCTAGAAAAAGAAAAGCAATCCAT  1858

Qy   961  GAGTACTTCTTGAGACTGTACGAATATTTCTGACGAATATTTCTGACAGAGGACTAGAAGAA  1020
          |||||||||||||||||||||||||||||||||||||
Db  1859  ----------TTCACTCAAACAAGATCTTCTTCAACACTGAAG  1904

Qy  1021  ACTTACCAGGAGCCCAGAGAGTCTGCCTCCTGAGGAGAGATGTGCCAGAAATGTCACTCGGCT  1080
          |||  || |||||
Db  1905  AATAAGTCAGTGACCCCAGAAGACGGGAAGCATGGCATGGCTGGA  1942

Qy  1081  TCTACGAAAGCAGCTGCGGAGGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGC  1140
          ||  ||||||
Db        ...
```

RESULT 11
AAD37242
ID   AAD37242 standard; DNA; 3446 BP.
XX
AC   AAD37242;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3447.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
XX
PN   WO200183695-A2.
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 53-54; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3447 containing nucleotides 1-1992 (N-terminus,
CC   hinge H1 and rods R1, R2 and R3), 8749-10227 (rod R24, hinge H4 and CR
CC   domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3446 BP; 1074 A; 766 C; 787 G; 819 T; 0 U; 0 Other;

Query Match   77.5%;   Score 1629;   DB 6;   Length 3446;
Best Local Similarity 87.9%;   Pred. No. 0;
Matches 1847;   Conservative 0;   Mismatches 190;   Indels 64;   Gaps 4;

```
Qy     1  GAGCTATGCCTACACACCAGGCTGCTTATGTCCACCACCTTCGACCCCTACACGGAGCCCATT  60
```

ID	ABK81999 standard; DNA; 5462 BP.	
XX		
AC	ABK81999;	
XX		
DT	13-AUG-2002 (first entry)	
XX		
DE	DNA encoding mini-dystrophin protein deltaR2-R21+H3.	
XX		
KW	Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;	
KW	Duchenne's muscular dystrophy; DMD; dystrophin; ds.	
XX		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
PN	WO20229056-A2.	
XX		
PD	11-APR-2002.	
XX		
PF	04-OCT-2001; 2001WO-US031126.	
XX		
PR	06-OCT-2000; 2000US-0238848P.	
XX		
PA	(UNMI) UNIV MICHIGAN.	
XX		
PI	Chamberlain JS, Harper SQ;	
XX		
DR	WPI; 2002-435334/46.	
XX		
PT	A composition for preparing therapeutic drugs, has a mini-dystrophin	
PT	peptide comprising a specific number of spectrin-like repeat domains, or	
PT	a nucleic acid sequence encoding the mini-dystrophin peptide.	
XX		
PS	Disclosure; Fig 14; 145pp; English.	
XX		
CC	The invention describes a composition comprising a mini-dystrophin	
CC	peptide comprising a spectrin-like repeat domain, where the domain	
CC	comprises n spectrin-like repeats, and contains no more than n spectrin-	
CC	like repeats, where n is an even number between 4-24, or a nucleic acid	
CC	encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the	
CC	polynucleotide encoding it is useful as a medicament, for preparing a	
CC	drug for therapeutic application and in the preparation of a composition	
CC	for treatment of muscle disease, e.g. Duchenne's muscular dystrophy	
CC	(DMD). This sequence represents a mini-dystrophin sequence of the	
CC	invention	
XX		
SQ	Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 U; 0 Other;	

```
Query Match             77.5%;  Score 1629;  DB 6;  Length 5462;
Best Local Similarity   85.4%;  Pred. No. 0;
Matches 1911;  Conservative    0;  Mismatches  190;  Indels  138;  Gaps   3;

Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy     61 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGAG 1218

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACACAGCTTTAGAGAGAAGTATTATCGTGGCTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACACAGCTTTAGAGAGAAGTATTATCGTGGCTCTTTC 1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGACACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGACACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy    241 CCAGTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGGCCGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTCATCATCATGAGGGGTACATGATGATTGACACGCCCATCAGGGGCCGGGTTGG 1398

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 13
ABK81997
ID   ABK81997 standard; DNA; 5417 BP.
XX
AC   ABK81997;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaR4-R23.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.

WO20029056-A2.

11-APR-2002.

04-OCT-2001; 2001WO-US031126.

06-OCT-2000; 2000US-0238848P.

(UNMI) UNIV MICHIGAN.

Chamberlain JS, Harper SQ;

WPI; 2002-435334/46.

A composition for preparing therapeutic drugs, has a mini-dystrophin peptide comprising a specific number of spectrin-like repeat domains, or a nucleic acid sequence encoding the mini-dystrophin peptide.

Disclosure; Fig 12; 145pp; English.

The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises n spectrin-like repeats, and contains no more than n spectrin-like repeats, where n is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament, for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention

Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 U; 0 Other;

Query Match   76.7%; Score 1611.6; DB 6; Length 5417;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1872; Conservative   0; Mismatches 229; Indels   93; Gaps   3;

```
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 46-47; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;

Query Match          76.2%;  Score 1602;  DB 6;  Length 3999;
Best Local Similarity 81.1%;  Pred. No. 0;
Matches 2101;    Conservative    0;  Mismatches   0;  Indels   489;  Gaps     1;

Qy       1  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCGCCATT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCGCCATT  959

Qy      61  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGGAGAG  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     960  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGGAGAG 1019

Qy     121  TGAAGTAAACCTGGACCGTATCAAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1020  TGAAGTAAACCTGGACCGTATCAAACACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079

Qy     181  TGCTGAGGACACAATGCAAGCACACAAGGAGAGATTTCTAATGATGTGGAGTGGTGAAAGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1080  TGCTGAGGACACACAATGCAAGCACACAAGGAGAGATTTCTAATGATGTGGAGTGGTGAAAGA 1139

Qy     241  CCAGTTCATCATCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGGTTGG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1140  CCAGTTCATCATCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGGTTGG 1199

Qy     301  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1200  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259

Qy     361  AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1260  AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy     421  TAGCATGGAAAAACAAAGCAATTTACATGAGTTTTAATGGATCTCCAGAATCAGAAACT  480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1320  TAGCATGGAAAAACAAAGCAATTTACATGAGTTTTAATGGATCTCCAGAATCAGAAACT 1379

Qy     481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAATGAGGAGGAGA  540
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1380  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAATGAGGAGGAGA 1439

Qy     541  GCCTCCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACAACATAAGGTGCTTCA  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1440  GCCTCCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACAACATAAGGTGCTTCA 1499

Qy     601  AGAAGATCTAGAACAGAACAAGAACAGTCAGGGTCAATTCTCCACTCACATGGTGGTAGT  660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1500  AGAAGATCTAGAACAGAACAAGAACAGTCAGGGTCAATTCTCCACTCACATGGTGGTAGT 1559

Qy     661  TGATGAATCTAGTGGAGATCAGCAACTGCTCGTGCTTTGGAAGAACAACTTAAGGTATTGGG  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1560  TGATGAATCTAGTGGAGATCAGCAACTGCTCGTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619
```

```
Qy    1528  TCTTTCCACGTCGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTA  1587
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2719  TCTTTCCACGTCGTCCAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTA  2778

Qy    1588  CTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTA  1647
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2779  CTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTA  2838

Qy    1648  CCAGTCTTTAGCTGACCTGAATAATGTCAGATTCAGCTTATAGGACTGCCATGAAACT  1707
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2839  CCAGTCTTTAGCTGACCTGAATAATGTCAGATTCAGCTTATAGGACTGCCATGAAACT  2898

Qy    1708  CCGAAGACTGCGAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCGCATGTGATGC  1767
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2899  CCGAAGACTGCGAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCGCATGTGATGC  2958

Qy    1768  CTTGGACCAGCAGCACAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAA  1827
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2959  CTTGGACCAGCAGCACAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAA  3018

Qy    1828  TTGTTTGACCACTATTTATGACCGGCCTGGAGCAAGGACGAAGCACAACAATTGGTCAACGTCCC  1887
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3019  TTGTTTGACCACTATTTATGACCGGCCTGGAGCAAGGACGAAGCACAACAATTGGTCAACGTCCC  3078

Qy    1888  TCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACAGG  1947
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3079  TCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACAGG  3138

Qy    1948  GAGGATCCGTGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGA  2007
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3139  GAGGATCCGTGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGA  3198

Qy    2008  AGACAAGTACAGATACACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCACGCG  2067
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3199  AGACAAGTACAGATACACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCACGCG  3258

Qy    2068  CAGGCTGGGCCTCCTTCTGCGATGATTCTATCCAA  2101
            |||||||||||||||||||||||||||||||||||
Db    3259  CAGGCTGGGCCTCCTTCTGCGATGATTCTATCCAA  3292
```

```
RESULT 14
AD37234
DB   AAD37234 standard; DNA; 3999 BP.
     AAD37234;
     21-AUG-2002   (first entry)
     Human dystrophin minigene delta3990.

     Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
     adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
     Becker muscular dystrophy; ds.
     Homo sapiens.

     WO200183695-A2.
     08-NOV-2001.
     27-APR-2001; 2001WO-US013677.
     28-APR-2000; 2000US-0200777P.
     (XIAO/) XIAO X.
     Xiao X;
     WPI; 2002-049342/06.
     New dystrophin minigene for treating Duchenne or Becker muscular
     dystrophy comprises an N-terminal domain or modified N-terminal domain,
```

```
  721 AGATCGATGGCAAACATCTGTAGATGATGGACAGAGACCGCTGGGTTCTTTTACAAGA---  777
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1620 AGATCGATGGCAAACATCTGTAGATGATGGACAGAGACCGCTGGGTTCTTTTACAAGACCA 1679

  778 -------------------------------------------------------------  777

 1680 GCCTGACCTAGCTCCTGGACTGACCACTATTGAGCCTCCTCCTACTCAGACTGTTACTCT  1739

  778 -------------------------------------------------------------  777

 1740 GGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC  1799

  778 -------------------------------------------------------------  777

 1800 CTTGATGTTGGAGGTACCTACTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGAAAA 1859

  778 -------------------------------------------------------------  777

 1860 GTTTCTTGCCTGGCTTACAGAAGCTGAAACACTGCCAATGTCCTACAGGATGCTACCCG  1919

  778 -------------------------------------------------------------  777

 1920 TAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCCAAGA 1979

  778 -------------------------------------------------------------  777

 1980 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACCTGGATGAAAACAGCCA   2039

  778 -------------------------------------------------------------  777

 2040 AAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAGTCCTGTTACAAGACGTTTGGA    2099

  778 -------------------------------------------------------------  777

 2100 TAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAGTCTCTCAACATTAGGTCCCATTT   2159

  778 ------CAGTTCTGACCAGTGGAAGCGTCTGACCTTTCTCTGCAGGAACTTCTGGTGTG   831
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2160 GGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTG 2219

  832 GCTACAGTCTGAAAGATGATGAATTAAGCCGGCAGCGCACCTATTGGAGCGCGACTTTCCAGC  891
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2220 GCTACAGTCTGAAAGATGATGAATTAAGCCGGCAGCGCACCTATTGGAGCGCGACTTTCCAGC 2279

  892 AGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACC   951
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2280 AGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACC  2339

  952 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAGG 1011
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2340 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAGG 2399

 1012 ACTCGGGCTTCTACGAGGAGCCAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT     1071
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2460 CACTCGGGCTTCTACGAGGAGCCAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT    2519

 1132 GCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTTCAGGAACTTCA  1191
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2520 GCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTTCAGGAACTTCA  2579

 1192 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTCAAGCTGATCAAGGGATCCTG 1251
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2580 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGATCAAGGGATCCTG      2639

 1252 GCAGCCCCGTGGGCGATCTCCTCCAGATCTCCTCCAAGATCACCTCAGAGAAAGTCAAGGC 1311
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2640 GCAGCCCCGTGGGCGATCTCCTCCAGATCTCCTCCAAGATCACCTCAGAGAAAGTCAAGGC 2699

 1312 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 1371
```

```
 2700 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCG 2759
 1372 CCAGCTTACCACTTTGGGCATTCAGCTTCAGAGCTCTCAGGTTCCAGTCCGTGAGGACCGAGTCAGCGACTCTGGAAGACCT 1431
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2760 CCAGCTTACCACTTTGGGCATTCAGCTTCAGAGCTCTCAGGTTCCAGTCCGTGAGGACCT 2819
 1432 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGCGCAGCTCGATGA 1491
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2820 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGCGCAGCTCGATGA 2879
 1492 AGCCCACGGAGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTGTCCAGGGTCC 1551
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2880 AGCCCACGGAGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTGTCCAGGGTCC 2939
 1552 CTGGGAGAGCAGCCATCTCGCCAAACAAAGTGCCCTACTATCAACCACGAGACTCAAAC 1611
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2940 CTGGGAGAGCAGCCATCTCGCCAAACAAAGTGCCCTACTATCAACCACGAGACTCAAAC 2999
 1612 AACTTGCTGGACCATCCCAAAATGACAGAGCTTTACCAGTCTTTAGCTGACCTGAATAA 1671
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 3000 AACTTGCTGGACCATCCCAAAATGACAGAGCTTTACCAGTCTTTAGCTGACCTGAATAA 3059
 1672 TGTCAGATTCTCAGCTTATAGGACCTGTCAGCTGGATGGCCTTGGACTGCAGAAGGCCCTTTG 1731
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 3060 TGTCAGATTCTCAGCTTATAGGACCTGTCAGCTGGATGGCCTTGGACTGCAGAAGGCCCTTTG 3119
 1732 CTTGGATTCTTGAGCCTGTGCAGCTGGATGTGATGCCATGAAAACTCCGAAGACTGCAGCA 1791
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
 3120 CTTGGATTCTTGAGCCTGTGCAGCTGGATGTGATGCCATGAAAACTCCGAAGACTGCAGCA 3179
 1792 AAATGACCACGCCCATGAGCCACAACAATTGGTCAACGTCCCCTCCCGTGGATGATGTGTCTGAA 1851
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
 3180 AAATGACCACGCCCATGAGCCACAACAATTGGTCAACGTCCCCTCCCGTGGATGATGTGTCTGAA 3239
 1852 CCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCCTCCCGTGGATGATGTGTCTGAA 1911
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
 3240 CCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCCTCCCGTGGATGATGTGTCTGAA 3299
 1912 CTGGCTGCTGAAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCGTGTCCTGTCTTTTAA 1971
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
 3300 CTGGCTGCTGAAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCCGTGTCTTTTAA 3359
 1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGAAGACAAGATACAGATACCTTTTCAA 2031
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
 3360 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGAAGACAAGATACAGATACCTTTTCAA 3419
 2032 GCAAGTGGCAAGTTCAACACAGGATTTTGTGACCAGCGCAGGCTTGGGCCTTCCTCCTTCTGCATGA 2091
      |||||||||||||||||||||||||||||||||||||||||||||||||||||
 3420 GCAAGTGGCAAGTTCAACACAGGATTTTGTGACCAGCGCAGGCTTGGGCCTTCCTCCTTCTGCATGA 3479
 2092 TTCTATCCAA 2101
      ||||||||||
 3480 TTCTATCCAA 3489
```

RESULT 15
AAD37256
```
ID   AAD37256 standard; DNA; 4966 BP.
XX
AC   AAD37256;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
```

```
XX
XN   WO200183695-A2.
XX
XD   08-NOV-2001.
XX
XP   27-APR-2001; 2001WO-US013677.
XX
XR   28-APR-2000; 2000US-0200777P.
XX
XA   (XIAO/) XIAO X.
XX
XI   Xiao X;
XX
XR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
     gene.
XX
PS   Example 1; Page 59-60; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;

Query Match       76.2%;   Score 1602;   DB 6;   Length 4966;
Best Local Similarity   81.1%;   Pred. No. 0;
Matches 2101;   Conservative   0;   Mismatches   0;   Indels  489;   Gaps   1;
```

```
Qy      1 GAGCTATGCTACCACACAGGCTGCTTATGTCACCACCCTGACCCTCACACGGAGCCCATT     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCTACCACACAGGCTGCTTATGTCACCACCCTGACCCTCACACGGAGCCCATT   1716

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   1776

Qy    121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGAAGCTTAGAGAAGAAGCAGCT    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAAGAAGCTTAGAGAAGAAGCAGCT   1836

Qy    181 TGCTGAGGACACATTGCAAGCACAACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  1896

Qy    241 CCAGTTTCATACTTCATGAGGGTACATGATGGGTACACAGCCCATCAGGGCCGGGTTGG    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTTCATACTTCATGAGGGTACATGATGGGTACACAGCCCATCAGGGCCGGGTTGG   1956

Qy    301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA   2016

Qy    361 AACTGAAGTACAAGAGCAGAGATGAATGGGAATGGCCTCAGGTAGC    420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGAGATGAATGGGAATGGCCTCAGGTAGC   2076

Qy    421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGAATCTCCAGAATCAGAAACT    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGAATCTCCAGAATCAGAAACT   2136
```

```
Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAAGAGAACAAGGAAAAATGAGGAAGA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAAGAGAACAAGGAAAAATGAGGAAGA  2196

Qy    541 GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA   2256

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT   2316

Qy    661 TGATGAATCTAGTGGAGATCACGCAACTCGTGCTTGGAAGAACAACTTAAGGTATTGGG    720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TGATGAATCTAGTGGAGATCACGCAACTCGTGCTTGGAAGAACAACTTAAGGTATTGGG   2375

Qy    721 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCCTGGGTTCTTTTACAAGA------   777
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACGACCCGCTGGGTTCTTTTACAAGACCA  2436

Qy    778 ------------------------------------------------------------   777

Db   2437 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTTCCTACTACGACTGTTACTCT   2496

Qy    778 ------------------------------------------------------------   777

Db   2497 GGTGACACAACCTGTGTGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC   2556

Qy    778 ------------------------------------------------------------   777

Db   2557 CTTGATGTGTGGAGGTACCTACTCATAGATTACTGCGAACAGTTCCCCTGGACCTGGAAAA   2616

Qy    778 ------------------------------------------------------------   777

Db   2617 GTTTCTTGCCTGGCCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGTCTACCCG   2676

Qy    778 ------------------------------------------------------------   777

Db   2677 TAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCAAGA   2736

Qy    778 ------------------------------------------------------------   777

Db   2737 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACCTGGATGAAGAAACAGCCA   2796

Qy    778 ------------------------------------------------------------   777

Db   2797 AAAAATCCTGAGATCCTGGAAGGTTCCGATGATGAGCAGTCGTGTTACAAAGACGTTTGGA   2856

Qy    778 ------------------------------------------------------------   777

Db   2857 TAACATGAACTTCAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGGTCCCATTT   2916

Qy    778 ------------------------------------------------------------   777

Db   2917 GGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGTGTGTG   2976

Qy    832 GCTACAGCTGAAAGATGATGAATGAATTAAGCCGGCAGCCACCTATTGGAGGCGCACTTTCCAGC   891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2977 GCTACAGCTGAAAGATGATGAATGAATTAAGCCGGCAGCCACCTATTGGAGGCGCACTTTCCAGC  3036

Qy    892 AGTTCAGAAGCAGAGAACGATGTACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAACC    951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3037 AGTTCAGAAGCAGAGAACGATGTACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAACC   3096

Qy    952 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCCTTTGGAAGG    1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3097 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCCTTTGGAAGG   3156

Qy   1012 ACTAGAGAAACTTCTACCAGGAGCCCAGAGACGCTGCCTCCTGAGGAGGAGAGCCCAGAAATGT    1071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3157 ACTAGAGAAACTTCTACCAGGAGCCCAGAGACGCTGCCTCCTGAGGAGGAGAGCCCAGAAATGT  3216

Qy   1072 CACTCGGCTTCTACGAAGCAGCAGGCTGAGGAGTCAATACTGAGTGGAAAATTGAACCT    1131
```

Mon Apr 5 15:45:50 2004   us-09-845-416-12_copy_900_3000.rng

Job time : 574.141 secs

```
3217 CACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTCGGGAAAAATTGAACCT 3276
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1132 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCA 1191
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3277 GCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCA 3336

1192 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTG 1251
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3337 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTG 3396

1252 GCAGCCCGTGGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 1311
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3397 GCAGCCCGTGGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 3456

1312 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTCG 1371
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3457 ACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTCG 3516

1372 CCAGCTTACCACTTTGGGCATTCAGCTTCCACGTATAAACCTCAGCACTCTGGAAGACCT 1431
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3517 CCAGCTTACCACTTTGGGCATTCAGCTTCCACGTATAAACCTCAGCACTCTGGAAGACCT 3576

1432 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTTGAGGACCGAGTCAGGCAGCTGCATGA 1491
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3577 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTTGAGGACCGAGTCAGGCAGCTGCATGA 3636

1492 AGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 1551
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3637 AGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGTCCAGGGTCC 3696

1552 CTGGGAGAGAGCCATCTCCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 1611
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3697 CTGGGAGAGAGCCATCTCCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 3756

1612 AACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGCTGACCTGAATAA 1671
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3757 AACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGCTGACCTGAATAA 3816

1672 TGTCAGATTCTCAGCTTATAGGACTGCCATGAGAACTCCGAAGACTGCGAAGGCCCCTTG 1731
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3817 TGTCAGATTCTCAGCTTATAGGACTGCCATGAGAACTCCGAAGACTGCGAAGGCCCCTTG 3876

1732 CTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCAACCTCAAGCA 1791
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3877 CTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCAACCTCAAGCA 3936

1792 AAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTTATGACCG 1851
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3937 AAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTTATGACCG 3996

1852 CCTGGAGCAAGAGCACAACAATTTGTCAACGTCCCTCCGTGGATATGTGTCTGAA 1911
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
3997 CCTGGAGCAAGAGCACAACAATTTGTCAACGTCCCTCCGTGGATATGTGTCTGAA 4056

1912 CTGGCTGCTGAAATGTTTATGATACGGGACGAACAAGGGAGGATCCGTGTCCTGTCTTTTAA 1971
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4057 CTGGCTGCTGAAATGTTTATGATACGGGACGAACAAGGGAGGATCCGTGTCCTGTCTTTTAA 4116

1972 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACACATTTGAAGACACATTTTCAA 2031
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4117 AACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACACATTTGAAGACACATTTTCAA 4176

2032 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTTCCTTCTGCATGA 2091
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
4177 GCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTTCCTTCTGCATGA 4236

2092 TTCTATCCAA 2101
     ||||||||||
4237 TTCTATCCAA 4246
```

earch completed: April 4, 2004, 18:20:51

% nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

in on: April 4, 2004, 12:04:07 ; Search time 5477.46 Seconds
                                      (without alignments)
                                      16625.161 Million cell updates/sec

.tle: US-09-845-416-12_COPY_900_3000
erfect score: 2101
equence: 1 gagctatgctacacacagg........tttctgcatgtgattctataccaa 2101

coring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

searched: 3470272 seqs. 2167516995 residues

otal number of hits satisfying chosen parameters: 6940544

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :   GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
       10: gb_ro:*
       11: gb_sts:*
       12: gb_sy:*
       13: gb_un:*
       14: gb_vi:*
       15: em_ba:*
       16: em_fun:*
       17: em_hum:*
       18: em_in:*
       19: em_mu:*
       20: em_om:*
       21: em_ov:*
       22: em_pat:*
       23: em_ph:*
       24: em_pl:*
       25: em_ro:*
       26: em_sts:*
       27: em_un:*
       28: em_vi:*
       29: em_vi:*
       30: em_htg_hum:*
       31: em_htg_inv:*
       32: em_htg_other:*
       33: em_htg_mus:*
       34: em_htg_pln:*
       35: em_htg_rod:*
       36: em_htg_mam:*
       37: em_htg_vrt:*
       38: em_sy:*
       39: em_htgo_hum:*
       40: em_htgo_mus:*
       41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1760.4	83.8	5339	6	AX538620	AX538620 Sequence
2	1629	77.5	5462	6	AX538621	AX538621 Sequence
3	1611.6	76.7	5417	6	AX538619	AX538619 Sequence
4	1328.8	63.2	5952	6	AR304538	AR304538 Sequence
5	1328.8	63.2	5952	6	AX114289	AX114289 Sequence
6	1328.8	63.2	8689	6	AX538622	AX538622 Sequence
7	1328.8	63.2	11443	6	AX538624	AX538624 Sequence
8	1328.8	63.2	12057	6	AX538627	AX538627 Sequence
9	1328.8	63.2	12446	6	HSDMDR	X14298 Human mRNA
10	1328.8	63.2	13957	6	AX409637	AX409637 Sequence
11	1328.8	63.2	13957	6	AX538581	AX538581 Sequence
12	1328.8	63.2	13957	9	HUMDYS	M18533 Homo sapien
13	1317.8	62.7	13977	6	AR220819	AR220819 Sequence
14	1317	62.7	4402	6	E30219	E30219 Shortened d
15	1196.6	57.0	4402	6	E30220	E30220 Shortened d
16	1185.4	56.4	13887	4	AF070485	AF070485 Canis fam
17	1149.6	54.7	13815	6	AX306153	AX306153 Sequence
18	1149.6	54.7	13815	6	AX538582	AX538582 Sequence
19	1149.6	54.7	13815	10	MUSDYSA	M68859 Mouse dystr
20 c	1149.6	54.7	19307	6	AR093392	AR093392 Sequence
21 c	1149.6	54.7	19307	6	AR142592	AR142592 Sequence
22	922.2	43.9	13575	5	GGDYS	X13369 Chicken mRN
23	718.8	34.2	4075	6	E30221	E30221 Shortened d
24	715.6	34.1	2654	5	FSCDYSTRO	M37645 Torpedo cal
25	667.2	31.8	3747	6	E30218	E30218 Shortened d
26	629.8	30.0	3275	6	MUSDYS	M18025 Mouse dystr
27	613.6	29.2	3163	6	E30223	E30223 Shortened d
28	598.2	28.5	3521	5	AF339031	AF339031 Danio rer
29	555.6	26.4	10705	10	RNAJ2967	AJ002967 Rattus no
30	554.2	26.4	11096	6	AX538584	AX538584 Sequence
31	554.2	26.4	11096	10	MMY12229	Y12229 M.musculus
32	554	26.4	3161	10	MMGUTRPH	X83506 M.musculus
33	547.4	26.1	6045	6	A63605	A63605 Sequence 7
34	547.4	26.1	6045	6	AR281528	AR281528 Sequence
35	547.4	26.1	6059	6	AX107972	AX107972 Sequence
36	547.4	26.1	10302	6	AX538583	AX538583 Sequence
37	547.4	26.1	10302	9	HSMUPS	X69086 H.sapiens m
38	547.4	26.1	10320	6	A63607	A63607 Sequence 9
39	547.4	26.1	10320	6	AR281529	AR281529 Sequence
40	528.4	25.1	3073	10	AF195788	AF195788 Rattus no
41	528.4	25.1	3140	10	AF195787	AF195787 Rattus no
42	527.4	25.1	3499	9	HSU43519	U43519 Human dystr
43	527.4	25.1	5106	6	AX552248	AX552248 Sequence
44	514.6	24.5	1481	9	BC009242	BC009242 Homo sapi
45	505	24.0	2005	6	AX817318	AX817318 Sequence

ALIGNMENTS

```
RESULT 1
AX538620
LOCUS       AX538620                5339 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 40 from Patent WO0229056.
ACCESSION   AX538620
VERSION     AX538620.1  GI:25271166
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 40 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
```

```
FEATURES             Location/Qualifiers
source               1..5339
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

Query Match      83.8%;   Score 1760.4;   DB 6;   Length 5339;
Best Local Similarity   90.9%;   Pred. No. 0;
Matches 1910;   Conservative  0;   Mismatches 191;   Indels  15;   Gaps   2;
```

```
Qy        1  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCTACACGCGAGCCCATT    60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1099  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCTACACGCGAGCCCATT  1158

Qy       61  TCCTTCACAGCATTTGGAAGCTCCTGAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1159  TCCTTCACAGCATTTGGAAGCTCCTGAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  1218

Qy      121  TGAAGTAAACTTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1219  TGAAGTAAACTTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  1278

Qy      181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1279  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA  1338

Qy      241  CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCCGGGTTGG   300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1339  CCAGTTTCATACTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGGCCCGGGTTGG  1398

Qy      301  TAATAATTCTACAATTGGGAAGTAAGCTGAATTGGAACAAGAGAAAAATTATCAAGATGAAGA   360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1399  TAATAATTCTACAATTGGGAAGTAAGCTGAATTGGAACAAGAGAAAAATTATCAAGATGAAGA  1458

Qy      361  AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1459  AACTGAAGTACAAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  1518

Qy      421  TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAATGATCTTCCAGAATCAGAAACT   480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1519  TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAATGATCTTCCAGAATCAGAAACT  1578

Qy      481  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAACGAGGAAAATGGAGGAGA   540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1579  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAACGAGGAAAATGGAGGAGA  1638

Qy      541  GCCTCTTGGACCTGATCTTGAAGACTTGAAGACCTAAAACGCCAAGTACAACACAATAAGGTGCTTCA   600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1639  GCCTCTTGGACCTGATCTTGAAGACTTGAAGACCTAAAACGCCAAGTACAACACAATAAGGTGCTTCA  1698

Qy      601  AGAAGATCTAGAACAAGAGAACAAGAGTCAGGGTCAATTCTCACTCGTGGTGGTAGT   660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1699  AGAAGATCTAGAACAAGAGAACAAGAGTCAGGGTCAATTCTCACTCGTGGTGGTAGT  1758

Qy      661  TGATGA--------ATCTAGTGAGATCACGCAACTGCTGTTTGGAAGAACAACT   708
             ||||||        |||||||||||||||||||||||||||||||||||||||
Db     1759  CAGCCAAAAAATCTGGAGATCCTGAGAGGTGGAAGGTCCTGGTCCTGTCTGAAGAACAACT  1818

Qy      709  TAAGG---TATTGGGAATGACTGAGTGAAGTGGAAACATCTGTAGATGGCAGAGAACCGCTGGGT   765
             |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1819  TTTGGATAACTGGGAATGACTGAGTGAAGTGGAAACATCTGTAGATGGCAGAGAACCGCTGGGT  1878

Qy      766  TCTTTTACAGACAGTCCAGTCTGACCGTCGAAGCGTCTGCACCTTTCTTGCTCCTCGAGGAACTTCT   825
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1879  CCATTTGGAAGCCAGTCTGACCGTCGAAGCGTCTGCACCTTTCTTGCTCCTCGAGGAACTTCT  1938

Qy      826  GGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCCAGGCACCTATTGGAGGCGACTT   885
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1939  GGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCCAGGCACCTATTGGAGGCGACTT  1998

Qy      886  TCCAGCAGTTCAGAAGCAGAACGATGCACATAGGGCCTTCAAGAGAGGGAATTGAAAACTAA   945
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1999  TCCAGCAGTTCAGAAGCAGAACGATGCACATAGGGCCTTCAAGAGAGGGAATTGAAAACTAA  2058

Qy      946  AGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT  1005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2059  AGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTT  2118

Qy     1006  GGAAGGACTAGAGAAACTCTACCAGAGAGCCCAGAGAGCTGCCTCCTGAGAGGAGAGCCCA  1065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2119  GGAAGGACTAGAGAAACTCTACCAGAGAGCCCAGAGAGCTGCCTCCTGAGAGGAGAGCCCA  2178

Qy     1066  GAATGTCACTCGGCTTCTACGAAAGCAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATT  1125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2179  GAATGTCACTCGGCTTCTACGAAAGCAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATT  2238

Qy     1126  GAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGA  1185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2239  GAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGA  2298

Qy     1186  ACTTCAAGAGGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGG  1245
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2299  ACTTCAAGAGGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGG  2358

Qy     1246  ATCCTGGCAGCCCGTGGGCCGATCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGT  1305
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2359  ATCCTGGCAGCCCGTGGGCCGATCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGT  2418

Qy     1306  CAAGGCACTTGAGAGAAATTGCGCCTCTGAAAGAACGTGAGCAGTCAACTCAATGACCT  1365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2419  CAAGGCACTTGAGAGAAATTGCGCCTCTGAAAGAACGTGAGCAGTCAACTCAATGACCT  2478

Qy     1366  TGCTCCCAGTTACCACTTTGGGCACTTCTCGCAGGTGGCCGTGAGCAGTCAGCGCAGCT  1425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2479  TGCTCGCCAGTTACCACTTTGGGCACTTCTCGCAGGTGGCCGTGAGCAGTCAGCGCAGCT  2538

Qy     1426  AGACCTGAACACCAGATGAGAAGCTTCAGCATCCAGACACTTCTTCTTCCACGTCTTGTCCA  1485
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2539  AGACCTGAACACCAGATGAGAAGCTTCAGCATCCAGACACTTCTTCTTCCACGTCTTGTCCA  2598

Qy     1486  GCATGAAGCCCACAGGGACTTTGGTCCAGCAACAAAGTGCCCTTACTATATCAACCGAGAC  1545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2599  GCATGAAGCCCACAGGGACTTTGGTCCAGCAACAAAGTGCCCTTACTATATCAACCGAGAC  2658

Qy     1546  GGGTCCCTGGGAGAGAGCCATCTGCCATGGACCTATAGGACTCGAGAACTGCGAGAAGGC  1605
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2659  GGGTCCCTGGGAGAGAGCCATCTGCCATGGACCTATAGGACTCGAGAACTGCGAGAAGGC  2718

Qy     1606  TCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT  1665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2719  TCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCT  2778

Qy     1666  GAATAATGTCAGCTTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCGAGAAGGC  1725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2779  GAATAATGTCAGCTTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCGAGAAGGC  2838

Qy     1726  CCTTTGCTTGGATCTCTTGAGCGTGCCTGTCAGCTGCGATGTGATGCCTTGGACCTGCAACAACCT  1785
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2839  CCTTTGCTTGGATCTCTTGAGCGTGCCTGTCAGCTGCGATGTGATGCCTTGGACCTGCAACAACCT  2898

Qy     1786  CAAGCAAAATGACCAGCCCATGGCAATCTGTAGATGGACAGAGAACCGCTGGGT  1845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2899  CAAGCAAAATGACCAGCCCATGGCAATCTGTAGATGGACAGAGAACCGCTGGGT  2958

Qy     1846  TGACCGCCTGGAGCAAGAGCAACAGACACAACATTGGTCAACGTCCCTCTGCGATATGTG  1905
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2959  TGACCGCCTGGAGCAAGAGCAACAGACACAACATTGGTCAACGTCCCTCTGCGATATGTG  3018

Qy     1906  TCTGAACTGGCTGCTGAAAGTTTATGATACGGACGAACAGGAGGATCCGTGTCCTGTC  1965
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3019  TCTGAACTGGCTGCTGAAAGTTTATGATACGGACGAACAGGAGGATCCGTGTCCTGTC  3078

Qy     1966  TTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGACATACTACACCT  2025
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
b   3079  TTTTAAAACTGGCATCATTCCCTGTGTAAAGCATCATTGGAAGACAAGTACAGATACCT 3138
y   2026  TTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGGCCTGGGCCTCCTTCT 2085
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3139  TTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGGCCTGGGCCTCCTTCT 3198

y   2086  GCATGATTCTATCCAA 2101
          ||||||||||||||||
b   3199  GCATGATTCTATCCAA 3214
```

ESULT 2
AX538621

LOCUS   AX538621   5462 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 41 from Patent WO0229056.
ACCESSION   AX538621
VERSION   AX538621.1 GI:25271168
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE   Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 41 11-APR-2002;
             THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
     source   1..5462
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

Query Match   77.5%;   Score 1629;   DB 6;   Length 5462;
Best Local Similarity   85.4%;   Pred. No. 0;
Matches 1911; Conservative   0; Mismatches 190; Indels 138; Gaps   3;

```
y      1  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCACCGACCCTACACCGAGCCCATT   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1099  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCACCGACCCTACACCGAGCCCATT 1158

y     61  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1159  TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1218

y    121  TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAGAAGTATTATCGTGCTTCTTTC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1219  TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAGAAGTATTATCGTGCTTCTTTC 1278

y    181  TGCTGAGGACACATTGCAAGCTGCAAGGCACACAGGAGAATTTCTAATGGTGGTGAAAGA  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1279  TGCTGAGGACACATTGCAAGCTGCAAGGCACACAGGAGAATTTCTAATGGTGGTGAAAGA 1338

y    241  CCAGTTTCATACTCATGAGGGGTACAGCGCATCATGATGATTTGACAGCCCATCAGGGGCCGGGTTGG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1339  CCAGTTTCATATCATGAGGGGTACAGCGCATCATGATGATTTGACAGCCCATCAGGGGCCGGGTTGG 1398

y    301  TAATATTCTACAATTGGGAAGTAGCTAGCTCTGATTGGAACAGGAGGAAAATTATCAGGAAGATGAAGA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1399  TAATATTCTACAATTGGGAAGTAGCTAGCTCTGATTGGAACAGGAGGAAAATTATCAGGAAGATGAAGA 1458

y    361  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAAATGCCTCAGGGTAGC  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1459  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGAAATGCCTCAGGGTAGC 1518

y    421  TAGCATGGAAAAACAAAGCAATTTACATAGAGTT------------------------  454
          |||||||||||||||||||||||||||||
b   1519  TAGCATGGAAAAACAAAGCAATTTACATGCTCCTGGACTGACCACTATTGGAGCCTCTCC 1578

       -----------------------------------------  454

b   1579  TACTCAGACTGTTACTCTGGTGACACACCTGTGTTACTAAGGAAACTGCCATCTCCAA 1638
```

```
b    455  -------------------------------------------------TTAATGGATCTCCAGAATCAGAA  477
y   1639  ACTAGAAATGCCATCTCTTCCTTGATGTGGAGCATAGATTACTGCAACAGTTCCCCTGGA 1698
          ||||||||||||||||||||||||||||||||||||
b   1698
```

```
y    478  ACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGAAAATGGAGGA  537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1699  CCTGAAAAGTTTCTTGCCTGGCTTACAGAGACCAGCTGAACACGTCCATGTCTACAGGA 1758

y    538  AGAGCCTCTTGGACCTGATCTTGAAGACTACCAAGCACATAAGGTGCT  597
          ||||||||||||||||||||||||||||||||||||||||||
b   1759  TGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACA 1818

y    598  TCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGTGGTGGT  657
          |||||||||||||||||||||||||||||||||||||||||||||
b   1819  ATGGCAAGACCTCCAAGTGAAATTGAAGTCTCACACAGATGTTATCACAACCTGGATGA 1878

y    658  AGTTGATGA------------ATCTAGTGGAGATCAGTGAGATGGGCAAACATCGTAGATGGACAGAGACACCGCTG  705
          |||||||                   ||||||||||||||||||||||||||
b   1879  AAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAG 1938

y    706  ACTTAAGG---TATTGGGATCGATGGGCAAACATCGTGTAGATGGACAGAGACACCGCTG  762
          |||||||                ||||||||||||||||||||||||||||||||
b   1939  ACGTTTGGATAACATGAACATTCAAGTGAGTGAACTTGGAAAAAGTCTCAACATTAG 1998

y    763  GGTTCTTTTACAGACAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACT  822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1999  GTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACT 2058

y    823  TCTGGTGTGCTACAGCTGAAAAGATGAATGAATTAAGCCGGCAGGCACCTATTGGAGCGA  882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2059  TCTGGTGTGCTACAGCTGAAAAGATGAATGAATTAAGCCGGCAGGCACCTATTGGAGCGA 2118

y    883  CTTTCCAGCAGTTCAGAGCAGCAGACGATGTACATAGGGGAATTGAAAAC  942
          ||||||||||||||||||||||||||||||||||||||||||||||||||
b   2119  CTTTCCAGCAGTTCAGAGCAGCAGACGATGTACATAGGGGAATTGAAAAC 2178

y    943  TAAAGAACCTGTAATCATGATGAGTACTCTGAGACTGTACGAATATTTCTGACAGAGCAGCC 1002
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2179  TAAAGAACCTGTAATCATGATGAGTACTCTGAGACTGTACGAATATTTCTGACAGAGCAGCC 2238

y   1003  TTTGGAAGGACTAGAGAGAACTCTACCAGGAGCTGCCTCCTGAGGACGAGAGC 1062
          ||||||||||||||||||||||||||||||||||||||||||||||||
b   2239  TTTGGAAGGACTAGAGAGAACTCTACCAGGAGCTGCCTCCTGAGGACGAGAGC 2298

y   1063  CCAGAATGTCACTCGGCTTCTACCGGGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAA 1122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2299  CCAGAAATGTCACTCGGCTTCTACCGGGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAA 2358

y   1123  ATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATGATGAAGACTCCA 1182
          ||||||||||||||||||||||||||||||||||||||||||||||||
b   2359  ATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAAGACTCCA 2418

y   1183  GGAACTTCAAGAGGCCACTTCGCGCCCTCAAGCTGCGCCAAGCTGAGTGATCAA 1242
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2419  GGAACTTCAAGAGGCCACTTCGCGCCCTCAAGCTGCGCCAAGCTGAGTGATCAA 2478

y   1243  GGGATCCTGCAGCCCGTGGGCGGCGATCCTCCATTGACTCGTGCTTCCAAGATCACCTCGAGAA 1302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2479  GGGATCCTGCAGCCCGTGGGCGGCGATCCTCATTGACTCGTGCTCCAAGATCACCTCGAGAA 2538

y   1303  AGTCAAGGCACTTCGCAGGAATTGCGCCCTCGAAAGAGAACGTGCCACGTCAATGA 1362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2539  AGTCAAGGCACTTCGCAGGAATTGCGCCCTCGAAAGAGAACGTGCCACGTCAATGA 2598

y   1363  CCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCAGCTCTATAACCTCAGCACTCT 1422
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   2599  CCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCAGCTCTATAACCTCAGCACTCT 2658

y   1423  GGAAGACCTGAACAACCAGATGGAAGCTTCTCGCAGGTGCCGTCGAGGACCGAGTCAGGCA 1482
          ||||||||||||||||||||||||||||||||||||||||||||||||||
b   2659  GGAAGACCTGAACAACCAGATGGAAGCTTCTCGCAGGTGCCGTCGGAGGACCGAGTCAGGCA 2718
```

```
Qy   1483 GCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGT 1542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2719 GCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTGT 2778

Qy     61 CCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCACTATATCAACCACGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2779 CCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCACTATATCAACCACGA 2838

Qy   1543 CCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCACTATATCAACCACGA 1602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2779 CCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCACTATATCAACCACGA 2838

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGAGCTTAGAGAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGACCGTTATCAAACAGAGCTTAGAGAGTATTATCGTGGCTTCTTTC 1278

Qy   1603 GACTCAAACAACTTGCTGGGACCATCCAAATGACACAGAGCTCTACCAGTCTTTAGCTGA 1662
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2839 GACTCAAACAACTTGCTGGGACCATCCAAATGACACAGAGCTCTACCAGTCTTTAGCTGA 2898

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   1663 CCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAGAAACTCCGAAGACTGCGAA 1722
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2899 CCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAGAAACTCCGAAGACTGCGAA 2958

Qy    241 CCAGTTTCATACTCCATGATCTGTCAGCTGCATGTGATTGACGCGCCATCAGGGCCGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCATGATCTGTCAGCTGCATGTGATTGACAGCGCCATCAGGGCCGGGTTGG 1398

Qy   1723 GGCCCTTGCTTGGATCTTTGAGCCTGCAGCTGCATGTGATTGACGCTTGGACCAGCACAA 1782
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2959 GGCCCTTGCTTGGATCTTTGAGCCTGCAGTCGCATGTGATTGACGCCTTGGACCAGCACAA 3018

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy   1783 CCTCAAGCAAATGACCAGCCAAGAGCACAACAATTGGTCAACGTCCCTCTCTCGCTGGATAT 1842
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3019 CCTCAAGCAAATGACCAGCCAAGAGCACAACAATTGGTCAACGTCCCTCTCTCGCTGGATAT 3078

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCCAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCTCCAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy   1843 TTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTCGCTGGATAT 1902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3079 TTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTCGCTGGATAT 3138

Qy    421 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATCGATCTCCAGAATCAGAAACT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATCGATCTCCAGAATCAGAAACT 1578

Qy   1903 GTGTCTGAACTGGCTGCTGAATGTTTATGAATACGGGACGAGGAGGATCCGTGTCCT 1962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3139 GTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGGACGAGGAGGATCCGTGTCCT 3198

Qy    481 GAAAGAGTTGAATGACTGGCTAACAAAAACACAGAACAGGAAAATGAGGAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACACAGAACAGGAAAATGAGGAAGA 1638

Qy   1963 GTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATA 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3199 GTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATA 3258

Qy    541 GCCTCTTGGACCTGATCTTGAAGACGCCAAGTACAACATAAGGTCTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACGCCAAGTACAACATAAGGTCTTCA 1698

Qy   2023 CCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTTCCT 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3259 CCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCCT 3318

Qy    601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCCACATGGTGGTGGTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCCACATGGTGTGTGGTAGT 1758

Qy   2083 TCTGCATGATTCTATCCAA 2101
          ||||||||||||||||||||
Db   3319 TCTGCATGATTCTATCCAA 3337

Qy    661 TGATGAATCTAGTGGAGATCGAGATCACGCACGCAACTGCTTGCTTGGAAGAACAACTTAAGGTATTGGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAGATCGAGATCACGCACGCAACTGCTTGCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy    721 AGATCGATGGGCAAACATCTGTAGATGGCACAGAAGACCGCTGGGTTCTTTTACAAGACAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy    781 TTCTGACCAGTGCAAGCGTCCGCACCTTTCTCTGCAGGAACTTCTGTGTGTGGCCTACAGCT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCCTTTTTAGTGCATGGCTTTCAGA 1938

Qy    841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGCGACTTTCAGCAGTTCAGAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAAGAATTCACACAACTGGCTTTAAAGATCAAAATGAAAATGTT 1998

Qy    901 GCAGAACGATGTACATAGGGCC--TTCAAGAGGGAATTGAAAAACTAAAGAACCTGTAATC 958
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGCCGATCTTACTGAAGAACAGGCGCCTTTTTAGTGCGTTCAGA 2058

Qy    959 ATGAGTACTCTTGAGACTGTACAGAATATTT--------CTGACAGAGCAGCCTTTGGA 1008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAACAAGATCTCTTTCAACACTGAAGATAAGTCAGTGAC 2118

Qy   1009 AGGACTAGAGAAACTCTACCAGAGCGCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAA 1068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAACGGAAGCGAGCATGGCTGGATAACTTGCCGGTGTTGGGATAATTTAGTCCAAAA 2178

Qy   1069 TGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGTCAATAC----------------- 1110
          ||||||||||||||||||||||||||||||||||
Qy      1 GAGCTGCATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
```

RESULT 3
AX538619

LOCUS   AX538619   5417 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 39 from Patent WO0229056.
ACCESSION   AX538619
VERSION   AX538619.1 GI:25271163
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
               artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE   Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 39 11-APR-2002;
               THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
     source   1..5417
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

Query Match   76.7%; Score 1611.6; DB 6; Length 5417;
Best Local Similarity   85.3%; Pred. No. 0;
Matches 1872; Conservative 0; Mismatches 229; Indels 93; Gaps 3;

Qy   1 GAGCTGCATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60

RESULT 4
AR304538

LOCUS　　　　　AR304538　　　　　　　　5952 bp　　　DNA　　　linear　　PAT 12-JUN-2003
DEFINITION　　Sequence 1 from patent US 6544786.
ACCESSION　　 AR304538
VERSION　　　　AR304538.1　　GI:31693691
KEYWORDS　　　.
SOURCE　　　　 Unknown.
　ORGANISM　　 Unknown.
　　　　　　　 Unclassified.
REFERENCE　　 1　(bases 1 to 5952)
　AUTHORS　　　Xiao,X. and Liu,P.X.
　TITLE　　　　Method and vector for producing and transferring trans-spliced
　　　　　　　 peptides
　JOURNAL　　　Patent: US 6544786-A 1 08-APR-2003;
FEATURES　　　　　　　　 Location/Qualifiers
　　　source　　　　　　　1..5952
　　　　　　　　　　　　　/organism="unknown"
　　　　　　　　　　　　　/mol_type="genomic DNA"
ORIGIN

Query Match　　　　63.2%;　Score 1328.8;　DB 6;　Length 5952;
Best Local Similarity　99.5%;　Pred. No. 0;
Matches 1333;　Conservative　0;　Mismatches　0;　Indels　7;　Gaps　0;

```
Qy    762 GGGTCTCTTTACAAGACAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC 821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3284 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC 3343

Qy    822 TTCTGGTGTGGCTACAGCTGAAAGATGAAAGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3344 TTCTGGTGTGGCTACAGCTGAAAGATGAAAGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 3403

Qy    882 ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3404 ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 3463

Qy    942 CTAAAGAACCTGTAATCAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTA 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3464 CTAAAGAACCTGTAATCAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTA 3523

Qy   1002 CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAG 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3524 CTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAG 3583

Qy   1062 CCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGAAAGCAGCTCAATACTGAGTGGGAAA 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3584 CCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGAAAGCAGCTCAATACTGAGTGGGAAA 3643

Qy   1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3644 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC 3703

Qy   1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCA 1241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3704 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCA 3763

Qy   1242 AGGGATCCTGGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCCAAGATCACCTCGAGA 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3764 AGGGATCCTGGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCCAAGATCACCTCGAGA 3823

Qy   1302 AAGTCAAGGCACTTGAGGAGAAATTGCGCCTTCTGAAAGAACGTGAGCCCACGTCAATG 1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3824 AAGTCAAGGCACTTGAGGAGAAATTGCGCCTTCTGAAAGAACGTGAGCCCACGTCAATG 3883

Qy   1362 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCCACCGTATAACCTCAGCACTC 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3884 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCCACCGTATAACCTCAGCACTC 3943

Qy   1422 TGGAAGACCTGAACACCAGATGAAGCTTCTGCGAGGTGGCCGTCGAGGAGATGCCCGAGAGGC 1481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

(Left column)

```
b    2179 ACTTGAAAAGAGTACAGCACAGATTCACAGGCTGTCACCACCACTCAGCCATCACTAAC 2238
y    1111 ----------------------------------------------------------- 1110
b    2239 ACAGACAACTGTAATGGAAAACAGTAACTACGGTGACCACAGCAGATCCTGGTAAA 2298
y    1111 ---TGAGTGGGAAAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGAC 1167
c    2299 GCATGCTCAAGAGAGGAACTTCCACCACCACCTCCCAAAGAGAGGCCAGATTACTGTGGA 2358
/    1168 CCTTGAAAGACTCCAGGACTTCAGGGACTTCAAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCGCCA 1227
c    2359 TCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCGCCA 2418
y    1228 AGTGGAGGTGATCAAGGGATCCTGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCA 1287
b    2419 AGCTGAGGTGATCAAGGGATCCTGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCA 2478
y    1288 AGATCACCTCGAGAAAGTCAAGGCACTTGCGCCTCTGAGAGAAATTGCGCCTCTGAAAGAACGT 1347
b    2479 AGATCACCTGAGAAAGTCAAGGCACTTGCGCCTCTGAGAGAAATTGCGCCTCTGAAAGAACGT 2538
y    1348 GAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCCTCCACCGTA 1407
b    2539 GAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCCTCCACCGTA 2598
y    1408 TAACCTCAGCACTCTGGAAGACTCTGAACACACAGATGGAAGCTTCTGCAGTGGCCGTCGA 1467
b    2599 TAACCTCAGCACTCTGGAAGACTCTGAACAACAGATGGAAGCTTCTGCAGTGGCCGTCGA 2658
y    1468 GGACCGAGTCAGGCAGCTGCATGAGAGCCCACAGGACTTTCCAGCATCTCAGCACTT 1527
b    2659 GGACCGAGTCAGGCAGCTGCATGAGAGCCCACAGGACTTTCCAGCATCTCAGCACTT 2718
y    1528 TCTTTCACGTCTGTCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTA 1587
b    2719 TCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTA 2778
y    1588 CTATATCAACCACGAGACTCAAACAACTGCTGGACCATCCCAAAATGACAGAGCTCTA 1647
b    2779 CTATATCAACCACGAGACTCAAACAACTGCTGGACCATCCCAAAATGACAGAGCTCTA 2838
y    1648 CCAGTCCTTAGCTGACCTGAATAATGTCAGCTTATAGGACTGCCATGAAACT 1707
b    2839 CCAGTCCTTAGCTGACCTGAATAATGTCAGCTTATAGGACTGCCATGAAACT 2898
y    1708 CCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGC 1767
b    2899 CCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGC 2958
y    1768 CTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAA 1827
b    3019 CTTGTTGACCATTATTATGACCGGCCTGGCTGGCTGTGATATCAGCCCACAATCCGCAGATTATTAA 3078
y    1828 TTGTTTGACCATTATTATGACCGGCCTGGCTGTGATATCAGCCCACAATCCGCAGATTATTAA 1887
b    3019 TTGTTTGACCATTATTATGACCGGCCTGGCTGTGATATCAGCCCACAATCCGCAGATTATTAA 3078
y    1888 TCTCTGCGTGGATATGTGCTGAACTGGCTGCTGCTGAATGTTTATGATACGGACGAACAGG 1947
b    3079 TCTCTGCGTGTGATATGTGCTGAACTGGCTGCTGCTGAATGTTTATGATACGGACGAACAGG 3138
y    1948 GAGGATCCGTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGA 2007
b    3139 GAGGATCCGTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGA 3198
y    2008 AGACAAGCACAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCG 2067
b    3199 AGACAAGCACAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCG 3258
y    2068 CAGGCTGGGCCTCCTTCCTTCTGCATGATTCTATCCAA 2101
b    3259 CAGGCTGGGCCTCCTTCCTTCTGCATGATTCTATCCAA 3292
```

```
ORIGIN                      /note="S2 junction site"

                 Query Match          63.2%;      Score 1328.8;    DB 6;    Length 5952;
                 Best Local Similarity 99.5%;      Pred. No. 0;
                 Matches 1333;   Conservative   0;   Mismatches   0;   Indels   7;   Gaps   0;

Qy     762  GGGTTCTTTACAGACAGTTCTGACCAGTGGAAGCGTCTGCACTTTCTCTGCAGGAAC  821
            ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    3284  GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACTTTCTCTGCAGGAAC 3343

Qy     822  TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGCG  881
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3344  TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGCG 3403

Qy     882  ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA  941
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3404  ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 3463

Qy     942  CTAAAGAACCTGTAATCATGGATACTCTTGACACTGTACGAATATTCTGACAGAGCAGC  1001
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3464  CTAAAGAACCTGTAATCATGGATACTCTTGACACTGTACGAATATTCTGACAGAGCAGC 3523

Qy    1002  CTTTGGAAGGACTAGAGAAACTCTACCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAG  1061
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3524  CTTTGGAAGGACTAGAGAAACTCTACCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAG 3583

Qy    1062  CCCAGAATGTCACTCGGCTTCACGTCCCTCCTGCGTGGATA  1121
            ||||||||||||||||||||||||||||||||||||||||||
Db    3584  CCAGAATGTCACTCGGCTTCACGTCCCTCCTGCGTGGGAAA 3643

Qy    1122  AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC  1181
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3644  AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC 3703

Qy    1182  AGGAACTTCAAGAGGCACCGATGAGCTGACCTCAAGCTGCAGTGAGTGATCA  1241
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3704  AGGAACTTCAAGAGGCACCGATGAGCTGACCTCAAGCTGCAGTGAGTGATCA 3763

Qy    1242  AGGGATCCTGCAGCCGTGGGCCATTCGAGGAGAAATTGCGCCTCCTGAAAGAGAACGTGAGCCACGTCAATG  1301
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3764  AGGGATCCTGCAGCCGTGGGCCATTCGAGGAGAAATTGCGCCTCCTGAAAGAGAACGTGAGCCACGTCAATG 3823

Qy    1302  AAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCCTGAAAGAGAACGTGAGCCACGTCAATG  1361
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3824  AAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCCTGAAAGAGAACGTGAGCCACGTCAATG 3883

Qy    1362  ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCGAGCTCTCACCTATAACCTCAGCACTC  1421
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3884  ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCGAGCTCTCACCTATAACCTCAGCACTC 3943

Qy    1422  TGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCGTCGAGGACCGAGTCAGGC  1481
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3944  TGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCGTCGAGGACCGAGTCAGGC 4003

Qy    1482  AGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG  1541
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4004  AGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 4063

Qy    1542  TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG  1601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4064  TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 4123

Qy    1602  AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG  1661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4124  AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 4183

Qy    1662  ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA  1721
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4184  ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 4243

Qy    1722  AGGCCCTTTGCTTGGATCTCTTGGACATCTTGAGCCTGTCAGCTGCATGTGATGCCTGCA  1781
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    ...

Db    3944  TGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCGTCGAGGACCGAGTCAGGC 4003

Qy    1482  AGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG  1541
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4004  AGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTG 4063

Qy    1542  TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG  1601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4064  TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 4123

Qy    1602  AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG  1661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4124  AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 4183

Qy    1662  ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA  1721
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4184  ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 4243

Qy    1722  AGGCCCTTTGCTTGGATCTCTTGGACATCTTGAGCCTGTCAGCTGCATGTGATGCCTGCA  1781
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    1782  ACCTCAAGCAAAATGACCAGCCATGGAATCCTGCAGATTATTAATGTTGTTGACCACTA  1841
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4304  ACCTCAAGCAAAATGACCAGCCATGGAATCCTGCAGATTATTAATGTTGTTGACCACTA 4363

Qy    1842  TTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCACGTCCCTCCTGCGTGGATA  1901
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4364  TTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCACGTCCCTCCTGCGTGGATA 4423

Qy    1902  TGTGTCTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCC  1961
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4424  TGTGTCTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGTCC 4483

Qy    1962  TGTCTTTTTAAAACTGCAATCATTTCCCTGTGTGTAAAGCACATTTGGAAGCAAGCAAGTACAGAT  2021
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4484  TGTCTTTTTAAAACTGCAATCATTTCCCTGTGTGTAAAGCACATTTGGAAGCAAGCAAGTACAGAT 4543

Qy    2022  ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCTCC  2081
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4544  ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGCCTCC 4603

Qy    2082  TTCTGCATGATTCTATCCAA  2101
            ||||||||||||||||||||
Db    4604  TTCTGCATGATTCTATCCAA 4623
```

```
RESULT 5
AX114289
LOCUS       AX114289                 5952 bp    DNA        linear        PAT 11-MAY-2001
DEFINITION  Sequence 1 from Patent WO0129243.
ACCESSION   AX114289
VERSION     AX114289.1  GI:14031259
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Paul,X.L. and Xiao,X.
  TITLE     Method and vector for producing and transferring trans -spliced
            peptides
  JOURNAL   Patent: WO 0129243-A 1 26-APR-2001;
            DALHOUSIE UNIVERSITY (CA); UNIV. OF PITTSBURGH OF THE COMMONWEALTH
            SYSTEM OF HIGHER EDUCATION (US)
FEATURES             Location/Qualifiers
     source          1..5952
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    2897..2898
                     /note="S4 junction site"
     misc_feature    3198..3199
```

RESULT 6
AX538622

LOCUS   AX538622   8689 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION   AX538622
VERSION   AX538622.1 GI:25271171
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE   Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 42 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
     source   1..8689
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic"
ORIGIN

Query Match   63.2%; Score 1328.8; DB 6; Length 8689;
Best Local Similarity   99.5%; Pred. No. 0;
Matches 1333; Conservative   0; Mismatches   0; Indels   7; Gaps   0;

Qy   762 GGGTTCTTTACAGACAGTTCTGACCAGTGAAGCCTGCACCTTTCTCTGCAGGAAC 821
         |||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db 3330 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCCTGCACCTTTCTCTGCAGGAAC 3389

Qy   822 TTCTGGTGTGGCTACAGTCGAAAGATGAAGATGAATTAAGCCGGCAGCCACCTATTGGAGGCG 881
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3390 TTCTGGTGTGGCTACAGTCGAAAGATGAAGATGAATTAAGCCGGCAGCCACCTATTGGAGGCG 3449

Qy   882 ACTTTCCAGCAGTTCAGAGAAACTTCAGAGAAGCAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 941
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3450 ACTTTCCAGCAGTTCAGAGAAACTTCAGAGAAGCAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 3509

Qy   942 CTAAAGAACCTGTAATCATGATACTCCTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3510 CTAAAGAACCTGTAATCATGATAGTACTCCTGAGACTGTACGAATATTTCTGACAGAGCAGC 3569

Qy 1002 CTTTGGAAGGAAGGACTAGGAAAACTTCACCAGGAGCCCAGAGCCTGCCTCCTGAGGAGAG 1061
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3570 CTTTGGAAGGAAGGACTAGGAGAAAACTTCACCAGGAGCCCAGAGCCTGCCTCCTGAGGAGAG 3629

Qy 1062 CCCAGAATGTCACTGCGCTTGCTTCTACGAAAGCAGCAGGCTGCGAGGAGTCAATACTGAGTGGGAAA 1121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3630 CCCAGAATGTCACTGCGCTTGCTTCTACGAAAGCAGCAGGCTGCGAGGAGTCAATACTGAGTGGGAAA 3689

Qy 1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC 1181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3690 AATTGAACCTGCACTCCGCTGACTGGCAGAGAGAAATAGATGAGACCCTTGAAAGACTCC 3749

Qy 1182 AGGAACTTCAAGAGGCCACGACTGAGTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 1241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3750 AGGAACTTCAAGAGGCCACGACTGAGTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 3809

Qy 1242 AGGGATCCTGGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCCCAAGATCACCTCGAGA 1301
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3810 AGGGATCCTGGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCCCAAGATCACCTCGAGA 3869

Qy 1302 AAGTCAAGGACACTTCGAGAGGAAATTGCGCCCTCGAAAGAACGTGAGCCACGTCAATG 1361
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3870 AAGTCAAGGACACTTCGAGAGAGAAATTGCGCCCTCGAAAGAACGTGAGCCACGTCAATG 3929

Qy 1362 ACCTTGCTCGCCAGCTTACCACTTGGCATCCAGCATTCAGCTTCACCGTATAACCTCAGCACTC 1421
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3930 ACCTTGCTCGCCAGCTTACCACTTGGCATCCAGCATTCAGCTTCACCGTATATAACCTCAGCACTC 3989

Qy 1422 TGGAAGACCTGAACACCAGCATGGAAGCTTCTGCAGGTGCCGTGGCCGTCGAGGACCGAGTCAGGC 1481
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3990 TGGAAGACCTGAACACCAGCATGGAAGCTTCTGCAGGTGCCGTGGCCGTCGAGGACCGAGTCAGGC 4049

Qy 1482 AGCTGCATGAAGCCCACAGGCACTTTGGTCAGCATCTCAGCACTTTCTTCTTCCAGTCTG 1541
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4050 AGCTGCATGAAGCCCACAGGACTTTTGGTCAGCATCTCAGCACTTTCTTCTTCCAGTCTG 4109

Qy 1542 TCCAGGTCCTGGGAGAGCCATCTGCCAAACAAAGTGCCTACTAATATCAACACG 1601
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4110 TCCAGGTCCTGGGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACACG 4169

Qy 1602 AGACTCAAACAACTGCTGTGGACCATCCAAAATGACAGAGCTTCTACCAGTCTTAGCTG 1661
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4170 AGACTCAAACAACTGCTGTGGACCATCCAAAATGACAGAGCTTCTACCAGTCTTAGCTG 4229

Qy 1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4230 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 4289

Qy 1722 AGGCCCTTTGCTTGGATCTCTTGAGCTCTGTCAGCTGCATGTGATGCCTTGGACCAGCACA 1781
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4290 AGGCCCTTTGCTTGGATCTCTTGAGCTCTGTCAGCTGCATGTGATGCCTTGGACCAGCACA 4349

Qy 1782 ACCTCAAGCAAATGACCAGCCATGAATATCCTGACGAATATCCTGAGATTATTGTTGACCACTA 1841
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4350 ACCTCAAGCAAATGACCAGCCATGAATATCCTGACGAATATCCTGAGATTATTGTTGACCACTA 4409

Qy 1842 TTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAAGTTCCTCTGCGTGATA 1901
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4410 TTTATGACCGCCTGGAGCAAGAGCAACAACAATTGGTCAAGTTCCTCTGCGTGATA 4469

Qy 1902 TGTGTCTGAACTGGCTGCTGAATGTTATGATACGGACGAACAGGAGATCCGTGTCC 1961
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4470 TGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGAGATCCGTGTCC 4529

Qy 1962 TGTCTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTGGAAGACAAGTACAGAT 2021
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4530 TGTCTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTGGAAGACAAGTACAGAT 4589

Qy 2022 ACCTTTCAAGCAAGTGGCAAGTTCAACAGCAGC 2081
         |||||||||||||||||||||||||||||||||
Db 4590 ACCTTTCAAGCAAGTGGCAAGTTCAACAGCAGC 4649

Qy 2082 TTCTGCATGATTCTATCCAA 2101
         ||||||||||||||||||||
Db 4650 TTCTGCATGATTCTATCCAA 4669

RESULT 7
AX538624
LOCUS   AX538624   11443 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 44 from Patent WO0229056.
ACCESSION   AX538624
VERSION   AX538624.1 GI:25271175
KEYWORDS   .
SOURCE   synthetic construct
ORGANISM   synthetic construct
               artificial sequences.
REFERENCE   1
AUTHORS   Chamberlain,J.S. and Harper,S.Q.
TITLE   Mini-dystrophin nucleic acid and peptide sequences
JOURNAL   Patent: WO 0229056-A 44 11-APR-2002;
               THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
source   1..11443
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Synthetic"
ORIGIN

Query Match   63.2%; Score 1328.8; DB 6; Length 11443;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy   762 GGGTTCTTTTACAAGACAGTTCTGACCAGTTGGAAGCCGTCTGCACTTTCTCTGCAGGAAC 821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6084 GGTCCCATTTGGAGCCAGTTCTGACCAGTTGGAAGCCGTCTGCACTTTCTCTGCAGGAAC 6143

Qy   822 TTCTGTGTGCTACAGCTGAAAGATGAATAAGCCGGCAGGCACCTATTGGAGGCG 881
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6144 TTCTGTGTGCTACAGCTGAAAGATGAATGAATAAGCCGGCAGGCACCTATTGGAGGCG 6203

Qy   882 ACTTTCCAGCAGTCAGAAGCAGAACAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 941
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6204 ACTTTCCAGCAGTCAGAAGCAGAACAGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 6263

Qy   942 CTAAAGAACCTGTAATCATGAGTACTCTGACTGAGAATATTTCTGACAGAGCAGC 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6264 CTAAAGAACCTGTAATCATGAGTACTCTGACTGAGAATATTTCTGACAGAGCAGC 6323

Qy   1002 CTTTGGAAGGACTAGAGAAACTTCACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGGAGAG 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6324 CTTTGGAAGGACTAGAGAAACTTCACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGGAGAG 6383

Qy   1062 CCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6384 CCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 6443

Qy   1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTC 1181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6444 AATTGAACCTGCACTCCGCTGACTGGCAGAGAGAAAAAATAGATGAGACCCTTGAAAGACTCC 6503

Qy   1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 1241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6504 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 6563

Qy   1242 AGGGATCCTGCAGCCGTGGGCCGATCTCATTGACTCTTCCAAGATCACCTTCGAGA 1301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6564 AGGGATCCTGCAGCCGTGGGCCGATCTCATTGACTCTTCCAAGATCACCTTCGAGA 6623

Qy   1302 AAGTCAAGGCACTTCGAGGAGAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATG 1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6624 AAGTCAAGGCACTTCGAGGAGAGAAATTGCCCCTCTGAAAGAGAAGACGTGAGCCACGTCAATG 6683

Qy   1362 ACCTTGCTCGGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6684 ACCTTGCTCGGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 6743

Qy   1422 TGGAAGACCTGAACACCAGCATGGAAGCTTCTGCAGGTGGCCGTGGACCGAGTCAGGC 1481
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6744 TGGAAGACCTGAACACCAGCATGGAAGCTTCTGCAGGTGGCCGTGGACCGAGTCAGGC 6803

Qy   1482 AGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG 1541
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6804 AGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTG 6863

Qy   1542 TCCAGGGTCCCTGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCACG 1601
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6864 TCCAGGGTCCCTGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCACG 6923

Qy   1602 AGACTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6924 AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGCTG 6983

Qy   1662 ACCTGAATATGTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCGAGA 1721
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6984 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAAGACTGCAGA 7043

Qy   1722 AGGCCCTTTGCTTGGATCTTCTTGAGCTGTCAGCTGCATGTGATGCCTTGACCAGCACA 1781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7044 AGGCCCTTTGCTTGTTGGATCTTCTTGAGCTGTCAGCTGCATGTGATGCCTTGACCAGCACA 7103

Qy   1782 ACCTCAAGCAAAATGACCAGCCCCATGGATAATCCTGCAGATTATTAATTGTTTGACCACTA 1841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7104 ACCTCAAGCAAAATGACCAGCCCCATGGATAATCCTGCAGATTATTAATTGTTTGACCACTA 7163

Qy   1842 TTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAAGTCCCTCCTCTGCGTGGATA 1901
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7164 TTTATGACCGGCTGGAGCAAGAGCACAACAATTTGGTCAACGCCCCTCTCTGCGTGGATA 7223

Qy   1902 TGTGTCTGACTGGCTGCTGAATGTTTATGATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC 1961
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7224 TGTGTCTGACTGGCTGCTGAATGTTTATGATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCC 7283

Qy   1962 TGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAGAT 2021
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7284 TGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAGAT 7343

Qy   2022 ACCTTTCAGCAAGTGGCAAGTCAACAGGATTTGTGACCAGCCCAGGCTGGGCCTCC 2081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7344 ACCTTTCAGCAAGTGGCAAGTCAACAGGATTTGTGACCAGCCCAGGCTGGGCCTCC 7403

Qy   2082 TTCTGCATGATTCTATCCAA 2101
          ||||||||||||||||||||
Db   7404 TTCTGCATGATTCTATCCAA 7423

RESULT 8
AX538627
LOCUS   AX538627   12057 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 47 from Patent WO0229056.
ACCESSION   AX538627
VERSION   AX538627.1 GI:25271181
KEYWORDS   .
SOURCE   synthetic construct
ORGANISM   synthetic construct
               artificial sequences.
REFERENCE   1
AUTHORS   Chamberlain,J.S. and Harper,S.Q.
TITLE   Mini-dystrophin nucleic acid and peptide sequences
JOURNAL   Patent: WO 0229056-A 47 11-APR-2002;
               THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
source   1..12057
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Synthetic"
ORIGIN

Query Match   63.2%; Score 1328.8; DB 6; Length 12057;
Best Local Similarity 99.5%; Pred. No. 0;

```
Matches 1333;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;

Qy   1842  TTTATGACCGCCTGGAGACGAAGAGCACAACAATTGGTCAACGTCCCTCTCGCGTGGATA  1901
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9678  TTTATGACCGCCTGGAGACGAAGAGCACAACAATTGGTCAACGTCCCTCTCGCGTGGATA  9737

Qy   1902  TGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCGTGTCC   1961
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9738  TGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCGTGTCC   9797

Qy   1962  TGTCTTTTAAAACTGGCATCATTTCCCTGTGAAGCACATTTGGAAGCAAGCAAGTACAGAT  2021
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9798  TGTCTTTTAAAACTGGCATCATTTCCCTGTGAAGCACATTTGGAAGCAAGCAAGTACAGAT  9857

Qy   2022  ACCTTTCAAGCAAGTGCAAGTCAACAGGATTTGTGACCAGCAGCGCAGGCTGGGCTCC    2081
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9858  ACCTTTCAAGCAAGTGCAAGTCAACAGGATTTGTGACCAGCAGCGCAGGCTGGGCTCC    9917

Qy   2082  TTCTGCATGATTCTATCCAA  2101
           ||||||||||||||||||||
Db   9918  TTCTGCATGATTCTATCCAA  9937
```

RESULT 9

HSDMDR

```
LOCUS       HSDMDR                 12446 bp    mRNA        linear   PRI 12-SEP-1993
DEFINITION  Human mRNA for dystrophin.
ACCESSION   X14298
VERSION     X14298.1  GI:30845
KEYWORDS    Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 12446)
  AUTHORS   Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
            Davies,K.E.
  TITLE     Two human cDNA molecules coding for the Duchenne muscular dystrophy
            (DMD) locus are highly homologous
  JOURNAL   Nucleic Acids Res. 17 (13), 5391 (1989)
  MEDLINE   89345106
   PUBMED   2668885
REFERENCE   2  (bases 1 to 12446)
  AUTHORS   Rosenthal,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
            der DDr, Zentralinstitut fuer Molekularbiologie, Robert-Roessle
            Str.10, 1115 Berlin Buch, DDR
COMMENT     see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
            M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449,
            3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES             Location/Qualifiers
     source          1..12446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X chromosome, Xp21."
                     /tissue_type="muscle"
                     /dev_stage="fetal and adult."
     CDS             99..11156
                     /note="unnamed protein product; dystrophin (AA 1 - 3685)"
                     /codon_start=1
                     /protein_id="CAA32479.1"
                     /db_xref="GI:30846"
                     /db_xref="GOA:P11532"
                     /db_xref="SWISS-PROT:P11532"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIILHWQVKNVMKNIMAGLQPTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEBHFQLHHQMYSQ
                     QITVSLAQGYERTSSPRPRFKSYAYTQAAYVTTSDPTRSPFPSQHEAPEDKSFGSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH
                     QGRVGNILQLGSKLIGTGKLSEDEETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
```

```
Qy    762  GGGTTCTTTACAGACAGTTCTGACCAGTGGAAGGCGTCTGCACCTTTCTCTGCAGGAAC   821
                   |||   ||
Db   8598  GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCCGTCTGCACCTTTCTCTGCAGGAAC  8657

Qy    822  TTCTGGTGTGGCTACAGCTGAAAGATGATGAAGATGATGAAT                    881
           |||||||||||||||||||||||||||||||||||||||||
Db   8658  TTCTGGTGTGGCTACAGCTGAAAGATGATGAAGATGATGAAT                    8717

Qy    882  ACTTTCCAGCAGTTCAGAAGCAGAACAGATGACAGAACCAGTCCTCAGAGAGGAATTGAAAA  941
           ||||||||||||||||||||||||||||||||||||||||||
Db   8718  ACTTTCCAGCAGTTCAGAAGCAGAACAGATGACAGAACCAGTCCTCAGAGAGGAATTGAAAA  8777

Qy    942  CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGACTGTACGAATATTTCTGACAGCAGC  1001
           |||||||||||||||||||||||||||||||||||||||
Db   8778  CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGACTGTACGAATATTTCTGACAGCAGC  8837

Qy   1002  CTTTGGAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCCTCCTGAGGAGGAGAG  1061
           |||||||||||||||||||||||||||||||||||||||
Db   8838  CTTTGGAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCCTCCTGAGGAGGAGAG  8897

Qy   1062  CCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGTCAATACTGAGTGGGAAA    1121
           |||||||||||||||||||||||||||||||||||||||
Db   8898  CCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGTCAATACTGAGTGGGAAA    8957

Qy   1122  AATTGAACCTGCACTCCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC  1181
           |||||||||||||||||||||||||||||||||||||||
Db   8958  AATTGAACCTGCACTCCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC  9017

Qy   1182  AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCGCCAAGCTGAGGTGATCA  1241
           |||||||||||||||||||||||||||||||||||||||
Db   9018  AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCGCCAAGCTGAGGTGATCA  9077

Qy   1242  AGGGATCCTGGCAGCCCGTGGGCGCCGCATCTCCTCCAAGATCACCTTCGAGA         1301
           |||||||||||||||||||||||||||||||||||||||
Db   9078  AGGGATCCTGGCAGCCCGTGGGCGCCGCATCTCCTCCAAGATCACCTTCGAGA         9137

Qy   1302  AAGTCAAGGCACTTGAGAACGTTGCGCCTGAAAGAGAACGTGAGCCACGTCAATG       1361
           |||||||||||||||||||||||||||||||||||||||
Db   9138  AAGTCAAGGCACTTGAGAACGTTGCGCCTGAAAGAGAACGTGAGCCACGTCAATG       9197

Qy   1362  ACCTTGCTGCCAGTCCACTTTGGGGATTCGACTCTCACCGTATAACCTCAGCACTC      1421
           |||||||||||||||||||||||||||||||||||||||
Db   9198  ACCTTGCTGCCAGTCCACTTTGGGGATTCGACTCTCACCGTATAACCTCAGCACTC      9257

Qy   1422  TGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCCGTGCGAGGACCGAGTCAGGC  1481
           |||||||||||||||||||||||||||||||||||||||
Db   9258  TGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCCGTGCGAGGACCGAGTCAGGC  9317

Qy   1482  AGCTGCATGAAGCCACCAGGACTTGCTGGTGTCCAGCATCCAGCACTTTCTTTCCACGGCTG  1541
           |||||||||||||||||||||||||||||||||||||||
Db   9318  AGCTGCATGAAGCCACCAGGACTTTGCTGGTGTCCAGCATCCAGCACTTTCTTTCCACGGCTG  9377

Qy   1542  TCCAGGGTCCCTGGGAGAGAGAATTGCGCCCATCTCCCCAAACAAAGTGCCCTACTATATCAACCACG  1601
           |||||||||||||||||||||||||||||||||||||||
Db   9378  TCCAGGGTCCCTGGGAGAGAGAATTGCGCCCATCTCCCCAAACAAAGTGCCCTACTATATCAACCACG  9437

Qy   1602  AGACTTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGAGTCTTAGCTG          1661
           |||||||||||||||||||||||||||||||||||||||
Db   9438  AGACTTGCAAACAACTTGCTGGACCATCCCAAAATGACAGAGAGTCTTTAGCTG         9497

Qy   1662  ACCTGGAATAATGTCAGATTCTCAGCTTATATAGGACTGCCATGAAACTCCGAAGACTGCAGA  1721
           |||||||||||||||||||||||||||||||||||||||
Db   9498  ACCTGGAATAATGTCAGATTCTCAGCTTATATAGGACTGCCATGAAACTCCGAAGACTGCAGA  9557

Qy   1722  AGGCCCTTGCTTGGATCTTCTTGAGCCTGTGCAGCTGCATGTGATGCTTGAGCCTTGGACCAGCACA  1781
           |||||||||||||||||||||||||||||||||||||||
Db   9558  AGGCCCTTGCTTGGATCTTCTTGAGCCTGTGCAGCTGCATGTGATGCTTGAGCCTTGGACCAGCACA  9617

Qy   1782  ACCTCAAGCAAAATGACCAGCCCATGATCCTGCGAGTATCCTGCAGATTATTAATTGTTTGACCACTA  1841
           |||||||||||||||||||||||||||||||||||||||
Db   9618  ACCTCAAGCAAAATGACCAGCCCATGATCCTGCGAGTATCCTGCAGATTATTAATTGTTTGACCACTA  9677
```

```
Qy    942 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGAGC 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8668 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGAGC 8727

Qy   1002 CTTTGGAAGGACTAGAGAAACTTACCACGAGAGAGCTGCCTCCTGAGGAGAGAG 1061
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8728 CTTTGGAAGGACTAGAGAAACTTACCACGAGAGAGCTGCCTCCTGAGGAGAGAG 8787

Qy   1062 CCCAGAATGTCACTCGGCTTCTACGAAACAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8788 CCCAGAATGTCACTCGGCTTCTACGAAACAGGCTGAGGAGGTCAATACTGAGTGGGAAA 8847

Qy   1122 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8848 AATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC 8907

Qy   1182 AGGAACTTCAAGAGCTGCCACCGATGAGCTGGCCAAGCTGCGCCAAGCTGAGGTGATCA 1241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8908 AGGAACTTCAAGAGCTGCCACCGATGAGCTGGCCAAGCTGCGCCAAGCTGAGGTGATCA 8967

Qy   1242 AGGGATCCTGCAGCCCGTGGCAGCTTCCTCCAGATCACCTCCGAGA 1301
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   8968 AGGGATCCTGCAGCCCGTGGCAGCTTCCTCCAGATCACCTCCGAGA 9027

Qy   1302 AAGTCAAGGCACTTCGAGGAGAAATTGCGGCCTTCGAAGAGAACGTGAGCCACGTCAATG 1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9028 AAGTCAAGGCACTTCGAGGAGAAATTGCGGCCTTCGAAGAGAACGTGAGCCACGTCAATG 9087

Qy   1362 ACCTTGCTCGCCAGCTTACCACTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTCTG 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9088 ACCTTGCTCGCCAGCTTACCACTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTCTG 9147

Qy   1422 TGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 1481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9148 TGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGC 9207

Qy   1482 AGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTCTG 1541
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9208 AGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTCTG 9267

Qy   1542 TCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 1601
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9268 TCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 9327

Qy   1602 AGACTGCAAACAACTGCTGGGACCATCCCAAAATGACAGAGACTCTACCAGTCTTAGCTG 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9328 AGACTGCAAACAACTGCTGGGACCATCCCAAAATGACAGAGACTCTACCAGTCTTAGCTG 9387

Qy   1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCATGCTGCCATGAAACTCCGAAGACTGCAGA 1721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9388 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCATGCTGCCATGAAACTCCGAAGACTGCAGA 9447

Qy   1722 AGGCCCTTTGCTTGGATCTCTTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCGTGATTATTAATTGTTTGACCACTA 1781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9448 AGGCCCTTTGCTTGGATCTCTTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCGTGATTATTAATTGTTTGACCACTA 9507

Qy   1782 ACCTCAAGCAACAAAATGACCATGAACAGCACAACAATTGGTCAAGTGCCCTCCTCGCGTGGATA 1841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9508 ACCTCAAGCAACAAAATGACCATGAACAGCACAACAATTGGTCAAGTGCCCTCCTCGCGTGGATA 9567

Qy   1842 TTTATGACCGCCTGAGACAGTGAAGAGCACAACAATTGGTCAAGTGCCCTCCTCGCGTGGATA 1901
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9568 TTTATGACCGCCTGAGACAGTGAAGAGCACAACAATTGGTCAAGTGCCCTCCTCGCGTGGATA 9627

Qy   1902 TGTGTCTGGAACTGGCTGCTGAAGTGTTATGATGACAGGGACGAACAGGGAGGGATCCGTGTCC 1961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9628 TGTGTCTGGAACTGGCTGCTGAATGTTATGATGACAGGGACGAACAGGGAGGGATCCGTGTCC 9687

Qy   1962 TGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGCAAGTACAGAT 2021
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9688 TGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGCAAGTACAGAT 9747
```

ORIGIN

```
Qy    762 GGGTTCTTTACAAGACAGTTCTGACCAGTGGAAGCGTCTGACACTTTCTCTGCAGGAAC 821
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8488 GGGTTCTTTACAAGACAGTTCTGACCAGTGGAAGCGTCTGACACTTTCTCTGCAGGAAC 8547

Qy    822 TTCTGGTGTGTCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 881
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8548 TTCTGGTGTGTCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG 8607

Qy    882 ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8608 ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA 8667
```

```
    2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACACAGATTTTGTGACCAGCGCAGGCTGGGCCTCC 2081
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    9748 ACCTTTTCAAGCAAGTGGCAAGTTCAACACAGATTTTGTGACCAGCGCAGGCTGGGCCTCC 9807

    2082 TTCTGCATGATTCTATCCAA 2101
         ||||||||||||||||||||
    9808 TTCTGCATGATTCTATCCAA 9827
```

SULT 10
409637
```
CUS        AX409637
DEFINITION Sequence 2284 from Patent WO0229103.   13957 bp   DNA   linear   PAT 14-JUN-2002
ACCESSION  AX409637
VERSION    AX409637.1  GI:21442342
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS    Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE      Gene expression profiles in liver cancer
JOURNAL    Patent: WO 0229103-A 2284 11-APR-2002;
           GENE LOGIC INC (US)
FEATURES   Location/Qualifiers
source     1..13957
           /organism="Homo sapiens"
           /mol_type="unassigned DNA"
           /db_xref="taxon:9606"
           /note="EMBL/GenBank Accession No. M18533"
ORIGIN
```

```
Query Match            63.2%;  Score 1328.8;  DB 6;  Length 13957;
Best Local Similarity  99.5%;  Pred. No. 0;
Matches 1333;  Conservative 0;  Mismatches 7;  Indels 0;  Gaps 0;
```

```
Qy   762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGGAAGGCGTCTGCACCTTTCTCTGCGAGGAAC 821
         ||||||||                |||||||||||||||||||||||||||||||||||||||
Db  8598 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGGCGTCTGCACCTTTCTCTGCGAGGAAC 8657

Qy   822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAAGCCGGCAGGCACCATTGGAGGCG 881
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8658 TTCTGGTGTGGCTACAGCTGAAAGATGATGAAGCCGGCAGGCACCATTGGAGGCG 8717

Qy   882 ACTTTCCAGCAGTTCAGAGAGCAGAAACGATGTACATAGGGCCTTCAGAGGGAATTGAAAA 941
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8718 ACTTTTCCAGCAGTTCAGAGAGCAGAAACGATGTACATAGGGCCTTCAGAGGGAATTGAAAA 8777

Qy   942 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGGACTGTACAGAGCAGC 1001
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8778 CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGGACTGTACAGAGCAGC 8837

Qy  1002 CTTTGGAAGACTAGAGAAACTCTACCAGGAGGCCCAGAGCCCAGAGAGCTGCCCTCCTGAGGAGGAGAG 1061
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8838 CTTTGGAAGACTAGAGAAACTCTACCAGGAGGCCCAGAGCCCAGAGAGCTGCCCTCCTGAGGAGGAGAG 8897

Qy  1062 CCCAGAAGTGCACTCGGCTTCTACGGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8898 CCCAGAAGTGCACTCGGCTTCTACGGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 8957

Qy  1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC 1181
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8958 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCC 9017

Qy  1182 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 1241
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9018 AGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 9077

Qy  1242 AGGGATCCTGGCAGCCCGTGGGCGCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 1301
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9078 AGGGATCCTGGCAGCCCGTGGGCGCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 9137
```

```
Qy  1302 AAGTCAAGGCACTTCGAGGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCAGGTCAATG 1361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9138 AAGTCAAGGCACTTCGAGGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCAGGTCAATG 9197

Qy  1362 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCAGCTATAACCTCAGCACTC 1421
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9198 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCAGCTATAACCTCAGCACTC 9257

Qy  1422 TGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGCGAGGACCGAGTCAGGC 1481
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9258 TGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGCGAGGACCGAGTCAGGC 9317

Qy  1482 AGCTGCATGAAGCCCACAGGGACTTTGGTCGTCCAGCATCTCAGCACTTCTTCCACGTCTG 1541
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9318 AGCTGCATGAAGCCCACAGGGACTTTGGTCGTCCAGCATCTCAGCACTTCTTCCACGTCTG 9377

Qy  1542 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTATATCAACCACG 1601
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9378 TCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTATATCAACCACG 9437

Qy  1602 AGACTCAAACAACTTGCTGGACCATCCAAAATGACAGAGTCTCTACCAGTCTTTAGCTG 1661
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9438 AGACTCAAACAACTTGCTGGACCATCCAAAATGACAGAGTCTCTACCAGTCTTTAGCTG 9497

Qy  1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGTCCATGAAACTCCGAAGACTGCAGA 1721
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9498 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGTCCATGAAACTCCGAAGACTGCAGA 9557

Qy  1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGCAGCTGCATGTGATGCCTTGACCAGCACA 1781
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9558 AGGCCCTTTGCTTGGATCTCTTGAGCCTGCAGCTGCATGTGATGCCTTGACCAGCACA 9617

Qy  1782 ACCTCAAGCAAAATGACCAGCCAAAATGACCAGCGATCCTGCAGAATATCCTGCAGATTATTAATTGTTGACCACTA 1841
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9618 ACCTCAAGCAAAATGACCAGCCAAAATGACCAGCGATCCTGCAGAATATCCTGCAGATTATTAATTGTTGACCACTA 9677

Qy  1842 TTTATGACCGCCTGGAGCAAGAGCACACAACAATTGGTCAACGTCCCTCTCTGCGTGGATA 1901
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9678 TTTATGACCGCCTGGAGCAAGAGCACACAACAATTGGTCAACGTCCCTCTCTGCGTGGATA 9737

Qy  1902 TGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1961
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9738 TGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 9797

Qy  1962 TGTCTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTGGAAGACAAGCAAGTACAGAT 2021
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9798 TGTCTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTGGAAGACAAGCAAGTACAGAT 9857

Qy  2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCC 2081
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9858 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCC 9917

Qy  2082 TTCTGCATGATTCTATCCAA 2101
         ||||||||||||||||||||
Db  9918 TTCTGCATGATTCTATCCAA 9937
```

```
RESULT 11
AX538581
LOCUS       AX538581    13957 bp    DNA    linear    PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO0229056.
ACCESSION   AX538581
VERSION     AX538581.1  GI:25271086
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Chamberlain,J.S. and Harper,S.Q.
TITLE       Mini-dystrophin nucleic acid and peptide sequences
JOURNAL     Patent: WO 0229056-A 1 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
```

```
FEATURES          Location/Qualifiers
     source       1..13957
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"

ORIGIN

Query Match      63.2%;  Score 1328.8;  DB 6;  Length 13957;
Best Local Similarity  99.5%;  Pred. No. 0;
Matches 1333;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;
```

```
Qy    762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC  821
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8598 GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAAC  8657

Qy    822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG  881
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8658 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCG  8717

Qy    882 ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCTTCAGAGGGGAATTGAAAA   941
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8718 ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCTTCAGAGGGGAATTGAAAA   8777

Qy    942 CTAAAGAACCTGTAATCATGATGAGTACTTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8778 CTAAAGAACCTGTAATCATGATGAGTACTTCTTGAGACTGTACGAATATTTCTGACACAGCAGC 8837

Qy   1002 CTTTGGAAGGACTAGAGAAACTCTACCAGGACGCGCTGCCTCCTGAGGAGAGAG 1061
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8838 CTTTGGAAGGACTAGAGAAACTCTACCAGGACGCGCTGCCTCCTGAGGAGAGAG 8897

Qy   1062 CCCAGAATGTCACTCGGCTTCACTCGGCTTCTACGAAAGCAGCTGGAGGAGGTCAATACTGAGTGGGAAA 1121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8898 CCCAGAATGTCACTCGGCTTCACTCGGCTTCTACGAAAGCAGCTGGAGGAGGTCAATACTGAGTGGGAAA 8957

Qy   1122 AATTGAACCTGACACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC 1181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8958 AATTGAACCTGACACTCCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCC 9017

Qy   1182 AGGAACTTCAAGAGCCACGGACCGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 1241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9018 AGGAACTTCAAGAGGCCACGGACCGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCA 9077

Qy   1242 AGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 1301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9078 AGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 9137

Qy   1302 AAGTCAAGGCACTTCGAGGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAATG 1361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9138 AAGTCAAGGCACTTCGAGGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAATG 9197

Qy   1362 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTTCACGTATAACTTCAGCACTC 1421
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9198 ACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTTCACCGTATAACTTCAGCACTC 9257

Qy   1422 TGGAAGACCTGACACCAGATGAAGCTTCTGCAGGTGGCCGTCGAGGACGGAGTCCAGGC 1481
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9258 TGGAAGACCTGACACCAGATGAAGCTTCTGCAGGTGGCCGTCGAGGACGGAGTCCAGGC 9317

Qy   1482 AGCTGCATGAAGCCCACAGGGACTTTGTCCAGCATCTCAGCACTTTCTTTCTTTCCAGTCTG 1541
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9318 AGCTGCATGAAGCCCACAGGGACTTTGTCCAGCATCTCAGCACTTTCTTTCTTTCCAGTCTG 9377

Qy   1542 TCCAGGGGTCCCTGGGAGGAGAGCCATCTCCGCCAAAGAAGTGCCCTACTATATCAACCACG 1601
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9378 TCCAGGGGTCCCTGGGAGGAGAGCCATCTCCGCCAAACAAGTGCCCTACTATATCAACCACG 9437

Qy   1602 AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTTAGCTG 1661
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9438 AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTTAGCTG 9497

Qy   1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   9498 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 9557

Qy   1722 AGGCCCTTTGCTTGCTTGGATCGCTGGATCCTCGCTGTCTCAGCTGCATGTGATGCCTGGACCTCAGCACA 1781
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9558 AGGCCCTTTGCTTGCTTGGATCGCTGGATCCTCGCTGTCTCAGCTGCATGTGATGCCTGGACCTCAGCACA 9617

Qy   1782 ACCTCAAGCAAAATGACCAGGCCCATGGATCATCCTGCAGATTATCCTGCAGATTATTAATGTTTGACCACTA 1841
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9618 ACCTCAAGCAAAATGACCAGGCCCATGGATCATCCTGCAGATTATCCTGCAGATTATTAATGTTTGACCACTA 9677

Qy   1842 TTTATGACCGGCTGAGCAAGAGCACAACAACAATTGGTCAACGTCCCTCTCTGCTGGATA 1901
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9678 TTTATGACCGGCTGAGCAAGAGCACAACAACAATTGGTCAACGTCCCTCTCTGCTGGATA 9737

Qy   1902 TGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGGAGGATCCGTGTCC 1961
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9738 TGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGGAGGATCCGTGTCC 9797

Qy   1962 TGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 2021
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9798 TGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGAT 9857

Qy   2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGCCAGGCTGGGCTCC 2081
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9858 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGCCAGGCTGGGCCTC 9917

Qy   2082 TTCTGCATGATTCTATCCAA 2101
          ||||||||||||||||||||
Db   9918 TTCTGCATGATTCTATCCAA 9937
```

```
RESULT 12
HUMDYS
LOCUS       HUMDYS                 13957 bp    mRNA    linear   PRI 25-MAY-2000
DEFINITION  Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION   M18533 M17154 M18026 M20250
VERSION     M18533.1  GI:181856
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1699)
  AUTHORS   Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and
            Kunkel,L.M.
  TITLE     Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
            preliminary genomic organization of the DMD gene in normal and
            affected individuals
  JOURNAL   Cell 50 (3), 509-517 (1987)
  MEDLINE   87273512
   PUBMED   3607877
REFERENCE   2  (bases 1678 to 3830)
  AUTHORS   Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
  TITLE     Conservation of the Duchenne muscular dystrophy gene in mice and
            humans
  JOURNAL   Science 238 (4825), 347-350 (1987)
  MEDLINE   88018015
   PUBMED   3659917
REFERENCE   3  (bases 1 to 13957)
  AUTHORS   Koenig,M., Monaco,A.P. and Kunkel,L.M.
  TITLE     The complete sequence of dystrophin predicts a rod-shaped
            cytoskeletal protein
  JOURNAL   Cell 53 (2), 219-226 (1988)
  MEDLINE   88194521
   PUBMED   3282674
COMMENT     On May 25, 2000 this sequence version replaced gi:340693.
            Draft entry and computer-readable sequence kindly provided by
            M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
            determined by the size of the deleted DNA segment. Deletions found
            in different patients were from positions 302-2200, 473-1168,
            1691-1810, and 1169-3011.
FEATURES             Location/Qualifiers
     source          1..13957
```

```
gene
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="X"
            /map="Xp21.3-p21.1"
            /tissue_type="muscle"
            /dev_stage="fetus"
            /note="G00-119-850"
            1..13957
            /gene="DMD"
CDS         209..11266
            /gene="DMD"
            /codon_start=1
            /product="dystrophin"
            /protein_id="AAA53189.1"
            /db_xref="GI:181857"
            /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
            QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNVNNKALRVLQNNNVDLVNIGSTDIVDG
            NHKLTLGLIWNIILHWQVKVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
            TSWSDGLALNALIHSHRPDLFDWNSVVCQIGKELLEHAFNIARYQLGIEKLLDPEDV
            DTTYPDKKSILMYITSLFQVLPQQVSIAIQEVEMLPRPPKVTKEEHFQLHHQMHYSQ
            QITVSLAQGYERTSSPKPRFKSYAYTQAAYVTTSDTSPPFSQHLEAPEDKSFGSSL
            MESEVNLDRYQTALEEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMDLITAH
            QGRVGNILQLGSKLIGTGKLSEDDETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
            DLQONKLKELNDWLTKTEERTRKMEEEPLGPDVNNILKRQQVQGHKVLQEDLEQEQVRVN
            SLTHMVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWVLLQDILLKWQRLTE
            EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLQKLAVLKADLEKKKQSMGKLYSLK
            QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTTQPSLTQTTVWN
            ETVTVTTREQILVKCHAQEELPPPPQGKRQITVDSEIRKRLDVDITELHSWITRSEA
            VLQSPEFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSAQALVEQMVNEGVNA
            DSIKQASEQLNSWIEFCQLLSERINWLEYQNNIAFNQLQEOMTTAENWLKIQ
            PTTPSEPTAIKSQLKICKDEVNRLSGLQPQIERLKIQSIALKEKGQPMFLDADFVAF
            TNHFKQVFSDVQAREKELQCIFDTLPPMRYQETMSAIRTWVQSETKLSIPQLSVTDY
            EIMEQRLGELQALQSSLQEQSGLIYLSTTVKEMSKKAPSEISRKYQSEFBEIEGRWK
            KLSSQLVEHCQKLEEQMNRLRKIQNHIQTLKKWMAEYDVFLKEEWPALGDSEILKKQL
            KQCRLLVSDITIQPSLNSVNEGGQKIKNEABPEFASRLETELKELNTONDHMCQQVY
            ARKEALKGGLEKTVSLQKDLSEMHEWMTQAEBEYLERDFEYKTPDELQKAVEEMRKAK
            EBAQQKEAKVLLLTEBSVNSVIAQAPPVAQEALKKELETLTTNYQWLCTRLNGKCKTLE
            EVWACWHELLSYLEKANKWLNEVEFKLKTTENIPGGAEEISEVLDSLENLMRHSEDNP
            NQIRILAQTLTDGGVMDELINEBELETFNSRWRELHEEAVRQKLLEQSIQSAQETEKS
            LHLIQESLITFIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLEEMKKHNQGKEA
            AQRVLSQIDVAGKELKQDVSMKFRLFQESMKFKRLQESKMILDEVKMHLPALETKSVE
            QEVVQSQLNHCVNLYKSLSEVKSBVEMVIKTGRQIVQKKQTENPKELDERVTALKLHY
            NELGAKVTERKQQLEKCLKLLSRMQRKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
            VAWGKATQKEIRKQKELQRVHLKSITEVGEALKTVEDKLSLLNSNWIAVTSRAE
            EWLNLLLEYQKHMETFDQNVDHITKWIIQADTLLDESEKKPQQKEDVLKRLKAELND
            IRPKVDSTRDQAANLMANRGDHCRKLVEFOISELNHRPAAISHRIKTGKASIPLKELE
            QFNSDIQKLLEPLEAEIQQGVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQQRITDER
            KREEIKIKQQLLQTKHNALKDLRSQRRKKALEISHQWYQYKRQADDLLKCLDDIEKKL
            ASLPEPRDERRKIEIDRELQKKKELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWRE
            IESKPAQFRRLNFAQIHITVREETMVMTEDMLEISYVPSTYLTEITHVSQALLSVEQ
            LLNAPDLCAKDFEDLFFKQEESLKNIKDSLQQSSRIDIIHSKKTAALQSATPVERVKL
            QEALSOLDFQWEKVNKMYKDRQGFPDRSVEKWRRFHYDIKIFNQWLTEAEQFLRKTQI
            PENWEHAKYKWYLKELQDGIGQRQTVRTLNATGEEIIQQSSKTDASILQEKLGSLNL
            RWQEVCKQLSDRKKRLEEQKNIISEFQRDLNEFVLWLEEADNIASIPLEPGKEQQLKE
            KLEQVLLVBELPLRQGIKQLNETGGQNYLVSAPISPEEQDKLENLKQTNLQNIKVS
            RALPEKQGEIEAQIKDLGQLEKKLEDLEEQLNHLLLWLSPIRNQLEIYNQPNQEGPFD
            VQEETEIAVQAKQPDVEEILSKGQHLYEKPATGPVKRKLEDLSSEWKAVNRLLQELRA
            KQPDLAPGLTTIGASPVTVTLVTQPVTVKETAISKLEMPSSLMLEVPALADFNRAWT
            ELTDWLSLLDQVIKSQRVMVGDLEDINEMIIKQKATMQDLEQRRPQLEELITAAQNLK
            NKTSNQEARTIITDRIERIQNQWDEVQEHLQNRRQQLNEMLKDSTQWLEAKEEAEQVL
            GQARAKLESWKEGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSAD
            DTRKVHMITENINASWRSIHKRVSERBAALEETHRLLQQFPLDLEKFLAWLTEAETTA
            NVLQDATRKERLLEDSKGVKELMKQWODLQGEIEAHTDVYHNLDENSQKILRSLEGSD
            DAVLLQRRLDNMNFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELLVWLQLKDDEL
            SRQAFIGGDFPAVQKQNVDHRAPKRELKTKVEPVIMSTLETVRIFLTEQPLBGLEKLYQ
            EPRELPPEERAQNVTRLLRKQAEEVNTEWEKLNLHSADWQRKIDETLERLQELQRATD
            ELDLKLRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEIAPLKENVSHVNDLARQL
            TTLGIQLSPYNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVQGP
            WERAISPNKVPYYINHETQTTCHPQFRKRRELIKRYWRYTDRLEQEHNNLVNVPLCVD
            LCLDLLSLAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHNNLVNVPLCVD
            MCLNWLLNVDTGRTGRIVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRL
            GLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKFBIEAALLDRWRLEPQSM
```

ORIGIN

```
Query Match          63.2%;  Score 1328.8;  DB 9;  Length 13957;
Best Local Similarity 99.5%;  Pred. No. 0;
Matches 1333;  Conservative 0;  Mismatches 7;  Indels 0;  Gaps 0;

Qy    762  GGGTTCTTTTACAAGACAGTCTTGACCAGTGGAGCCTCTGACACCTTTCTCTGCAGGAAC  821
                 ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   8598  GGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCCACCTTTCTCTGCAGGAAC  8657

Qy    822  TTCTGGTGTGGCTACAGCTGAAAGATGAATTAAGCCGCAGGGCACCTATTGGAGGCG  881
                 ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   8658  TTCTGGTGTGGCTACAGCTGAAAGATGAATTAAGCCGGCAGGGCACCTATTGGAGGCG  8717

Qy    882  ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA  941
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8718  ACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAA  8777

Qy    942  CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC  1001
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8778  CTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC  8837

Qy   1002  CTTGGAAGACTAGAGAACTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGGAGAG  1061
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8838  CTTGGAAGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGGAGAG  8897

Qy   1062  CCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGTCAATACTGAGTGGGAAA  1121
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8898  CCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGTCAATACTGAGTGGGAAA  8957

Qy   1122  AATTGAACCTGCACTCCCGCTGGCACGAGAAAAATAGATGAGACCCTTGAAAGACTCC  1181
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8958  AATTGAACCTGCACTCCCGCTGGCACGAGAAAAATAGATGAGACCCTTGAAAGACTCC  9017

Qy   1182  AGGAACTTCAAGAGGCCACGGATGAGCTGAGCTGGCGCCAAGCTGAGGTGATCA  1241
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9018  AGGGAACTTCAAGAGGCCACGGATGAGCTGAGCTGGCGCCAAGCTGAGGTGATCA  9077

Qy   1242  AGGGATCCTGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACCTGAGGCCACGTCAATG  1301
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9078  AGGGGATCCTGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACCTGAGGCCACGTCAATG  9137

Qy   1302  AAGTCAAGGCACTTGGGGCATTGGGGCATTCAGCTTCACCGTATAACCTCAGCACTC  1361
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9138  AAGTCAAGGCACTTGGGGCATTGGGGCATTCAGCTTCACCGTATAACCTCAGCACTC  9197

Qy   1362  ACCTTGCTCGCCAGCTTACCACTTGGTCCAGCATCTCGCAGGTGGCCGTCGAGGACCGAGTCAGGC  1421
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9198  ACCTTGCTCGCCAGCTTACCACTTGGTCCAGCATCTCGCAGGTGGCCGTCGAGGACCGAGTCAGGC  1481

Qy   1422  TGGAAGACCTGAACACCAGATGGAAGCTTCGCAGGTGGCCGTCGAGGACCGAGTCAGGC  1481
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9258  TGGAAGACCTGAACACCAGATGGAAGCTTCGCAGGTGGCCGTCGAGGACCGAGTCAGGC  9317

Qy   1482  AGTTGCATGAAGCCCACAGGGACCTTGGTCCAGCATCTCGCAGCATCCAGCACTTCCCACGTCTG  1541
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9318  AGCTGCATGAAGCCCACAGGGACCTTGGTCCAGCATCTCGCAGCATCCAGCACTTCCCACGTCTG  9377

Qy   1542  TCCAGGGTCCCTGGGAGAGAGCCATCCCAAAACAAAGTGCCCAAAATCATCAACCACG  1601
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9378  TCCAGGGTCCCTGGGAGAGAGCCATCCCAAAACAAAGTGCCCAAAATCATCAACCACG  9437

Qy   1602  AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG  1661
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9438  AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTG  9497
```

```
Qy  1002 CTTTGGAAGACTAGAGAAACTCTACCAGGAGAGGCCGCCTCCTGAGGAGGAG 1061
         ||||||||||||||||||||||||||||||||||        |||||||||||||
Db  8838 CTTTGGAAGACTAGAGAAACTCTACCAGGAGAGCCCAGGAGAGCTGCCTCCTGAGGAGGAG 8897

Qy  1062 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGTGAGGAGGTCAATACTGAGTGGGAAA 1121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8898 CCCAGAATGTCACTCGGCTTCTACGAAAGCAGGTGAGGAGGTCAATACTGAGTGGGAAA 8957

Qy  1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAGAAAAATAGAGACCCTTGAAAGACTCC 1181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8958 AATTGAACCTGCACTCCGCTGACTGGCAGAGAGAAAAATAGAGACCCTTGAAAGACTCC 9017

Qy  1182 AGGAACTTCAAGAGGCCACGATGAGCTGACCTCAAGCTGCGCGCCAAGCTGAGGTGATCA 1241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9018 AGGAACTTCAAGAGGCCACGATGAGCTGACCTCAAGCTGCGCGCCAAGCTGAGGTGATCA 9077

Qy  1242 AGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 1301
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9078 AGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGA 9137

Qy  1302 AAGTCAAGGCACTTGAGGAGGAAAATGCGCCTCTGAAAGAAGAACGTGAGCCACGTCAATG 1361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9138 AAGTCAAGGCACTTGAGGAGGAAAATGCGCCTCTGAAAGAAGAACGTGAGCCACGTCAATG 9197

Qy  1362 ACCTTGCTGCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9198 ACCTTGCTGCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTC 9257

Qy  1422 TGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGC 1481
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9258 TGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAGGACCGAGTCAGGC 9317

Qy  1482 AGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTG 1541
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9318 AGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTG 9377

Qy  1542 TCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 1601
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9378 TCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACG 9437

Qy  1602 AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGCTG 1661
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9438 AGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGCTG 9497

Qy  1662 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 1721
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9498 ACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGA 9557

Qy  1722 AGGCCCTTGCTTGGATCTCTTGGAGCTGTCAGCTGCGTCAGTGTGATGCCTTGGACCAGCACA 1781
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9558 AGGCCCTTGCTTGGATCTCTTGGAGCTGTCAGCTGCGTCAGTGTGATGCCTTGGACCAGCACA 9617

Qy  1782 ACCTGAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 1841
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9618 ACCTGAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTA 9677

Qy  1842 TTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTGCGTGGATA 1901
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9678 TTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTGCGTGGATA 9737

Qy  1902 TGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 1961
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9738 TGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCC 9797

Qy  1962 TGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGACACATTGGAAGACAAGTACAGAT 2021
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9798 TGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGACACATTGGAAGACAAGTACAGAT 9857

Qy  2022 ACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTC 2080
         ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  9858 ACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGGCCCTC 9917

Qy  2081 TTCTGCATGATTCTATCCAA 2101
         ||||||||||||||||||||
Db  9918 TTCTGCATGATTCTATCCAA 9937
```

RESULT 13
AR220819

```
LOCUS       AR220819        13977 bp    DNA        linear    PAT  26-SEP-2002
DEFINITION  Sequence 60 from patent US 6426186.
ACCESSION   AR220819
VERSION     AR220819.1    GI:23327696
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 13977)
AUTHORS     Jones,K.A., Volkmuth,W. and Walker,M.G.
TITLE       Bone remodeling genes
JOURNAL     Patent: US 6426186-A 60 30-JUL-2002;
FEATURES    Location/Qualifiers
source      1..13977
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
```

Query Match 62.7%; Score 1317.8; DB 6; Length 13977;
Best Local Similarity 99.4%;
Matches 1333; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

```
Qy   762 GGGTTCTTTTACAAGACAGTTCTGACCAGTGGAAGCCGTCTGCACCTTTCTCTGCAGGAAC 821
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8598 GTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCCGTCTGCACCTTTCTCTGCAGGAAC 8657

Qy   822 TTCTCGGTGTGGCTACAGCTGAAAGATGAATGAATTAAGCCGGCACCTATTGGAGGCG 881
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8658 TTCTCGGTGTGGCTACAGCTGAAAGATGAATGAATTAAGCCGGCACCTATTGGAGGGCG 8717

Qy   882 ACTTTCCAGCAGTTCAGAAGCAGAAGCAGATGTACATAGGGCGTACATAGGGAATTGAAAA 941
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8718 ACTTTCCAGCAGTTCAGAAGCAGAAGCAGATGTACATAGGGCGTACATAGGGAATTGAAAA 8777

Qy   942 CTAAAGAACCTGTAATCATGAGTACTCTTGACACTGTACGACTGTACGAATATTTCTGACAGAGCAGC 1001
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8778 CTAAAGAACCTGTAATCATGAGTACTCTTGACACTGTACGACTGTACGAATATTTCTGACAGAGCAGC 8837

Qy  2081 CTTCGCATGATTCTATCCAA 2101
```

```
                    |||||||||||||||||||||||
           9918 CTTCTGCATGATGATTCTATCCAA 9938


RESULT 14
30219
LOCUS        E30219        4402 bp    DNA    linear    PAT 18-JUN-2001
DEFINITION   Shortened dystrophin.
ACCESSION    E30219
VERSION      E30219.1  GI:13017026
KEYWORDS     JP 1999318467-A/2.
SOURCE       unidentified
ORGANISM     unidentified
             unclassified.
REFERENCE    1  (bases 1 to 4402)
AUTHORS      Sinichi,T.
TITLE        Shortened dystrophin
JOURNAL      Patent: JP 1999318467-A 2 24-NOV-1999;
             SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY
COMMENT      OS   Unidentified
             PN   JP 1999318467-A/2
             PD   24-NOV-1999
             PF   08-MAY-1998 JP 1998142134
             PR
             PI   SINICHI TAKEDA
             PC   C12N15/09,A61K48/00,C12N15/00
             CC   Strandedness: Both;
             CC   Topology: Linear;
             FH   Key             Location/Qualifiers
             FT   source          1..4402
FEATURES             Location/Qualifiers
source               1..4402
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Query Match       62.7%;  Score 1317;  DB 6;  Length 4402;
Best Local Similarity   81.6%;  Pred. No. 0;
Matches 1714;  Conservative    0;  Mismatches   0;  Indels  387;  Gaps    1;

Qy        1 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT     60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1095 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   1154

Qy       61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG    120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1155 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   1214

Qy      121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGTGCTTCTTC    180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1215 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGTGCTTCTTTC   1274

Qy      181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCAATGATGTGGAAGTGGTGAAAGA    240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1275 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCAATGATGTGGAAGTGGTGAAAGA   1334

Qy      241 CCAGTTCATCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGGCCGGGTGG    300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1335 CCAGTTCATCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGGCCGGGTGG   1394

Qy      301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA    360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1395 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGATGAAGA   1454

Qy      361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC    420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1455 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   1514

Qy      421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT    480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1515 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   1574

Qy      481 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAGAAGAACAAGGAAAATGGAGGAAGA    540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1575 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAGAAGAACAAGGAAAATGGAGGAAGA   1634

Qy      541 GCCTCTTGGACCTGATCTTGAAGACTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA    600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1635 GCCTCTTGGACCTGATCTTGAAGACTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA   1694

Qy      601 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGTAGT    660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1695 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGTAGT   1754

Qy      661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGACAACTTAAGGTATTGGG    720
            ||||||||||||||||||||||||||||||||||||||||||||||||---------
Db     1755 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTA--------   1804

Qy      721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAG    780
            -------------------------------------------------
Db     1805 -------------------------------------------------------   1804

Qy      781 TTCTGACCAGTGGAAGGCTCTGCACCTTTCTCTGCAGGAGAACTTCTGTGTGTGGCTACAGCT    840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1805 -------------------------------------------------------   1804

Qy      841 GAAAGATGATGAATTAGCCGGAGGCACCTATTGGAGGCGACCTTTCCAGCAGTTCAGAA    900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1805 -------------------------------------------------------   1804

Qy      901 GCAGAACGATGTACATAGGGCGTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCAT    960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1805 -------------------------------------------------------   1804

Qy      961 GAGTACTCTTGAGACGTGTACGAATATTTCGACAGAGCAGCCTTTGGAAGGACTAGAGAA   1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1805 -------------------------------------------------------   1804

Qy     1021 ACTCTACCAGGAGCCCAGAGAGTCGCTCCTCTGAGGAGAGCCCAGAATGTCACTCGGCT   1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1805 -------------------------------------------------------   1804

Qy     1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGAAAATTGAACCTGCACTCCGC   1140
            ||||||||||||||||||||||||||||||||||  |||||||||||||||||||||
Db     1805 --------------------------AGGAGGTCAATACTGAGTGGAAAATTGAACCTGCACTCCGC   1847

Qy     1141 TGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCAGGAACTTCAAGAGGCCAC   1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1848 TGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCAGGAACTTCAAGAGGCCAC   1907

Qy     1201 GGATGAGCTGGACCTCAAGCTGCGCGCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGT   1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1908 GGATGAGCTGGACCTCAAGCTGCGCGCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGT   1967

Qy     1261 GGGCGATCTCCTCATTGACTCCTCTCTCCAAGATCACCTCCAAGGCACTTCGAGG   1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1968 GGGCGATCTCCTCATTGACTCCTCTCTCCAAGATCACCTCCAAGGCACTTCGAGG   2027

Qy     1321 AGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC   1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2028 AGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC   2087

Qy     1381 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCGAGACCTGAACACCAG   1440
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db     2088 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCGAGACCTGAACACCAG   2147

Qy     1441 ATGGAAGCTTCTGCAGGTGGCCGTGCGAGGACCCGATGCAGCTGCAGGTAGC   1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2148 ATGGAAGCTTCTGCAGGTGGCCGTGCGAGGACCCGATGCAGCTGCAGGTAGC   2207

Qy     1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAG   1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2208 GGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCGTCCAGGGTCCCTGGGAGAG   2267
```

```
Qy  1561 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 1620
         |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2268 AGCCATCCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 2327

Qy  1621 GGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2328 GGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 2387

Qy  1681 CTCAGCTTATAGGACTGCCATGAAAACTCCGAAGACTGCGAGAAGGCCCTTGCTTGGATCT 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2388 CTCAGCTTATAGGACTGCCATGAAAACTCCGAAGACTGCGAGAAGGCCCTTGCTTGGATCT 2447

Qy  1741 CTTGAGGCTGTCAGCTGCGTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2448 CTTGAGGCTGTCAGCTGCGTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCA 2507

Qy  1801 GCCCATGATGATCCTGCGAGATTATTAATTGTTGACCACCATTATTATGACTGGGAGCA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2508 GCCCATGATGATCCTGCGAGATTATTAATTGTTGACCACCATTATTATGACTGGGAGCA 2567

Qy  1861 AGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGGATATGTGTCTGAACTGGCTGCT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2568 AGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGGATATGTGTCTGAACTGGCTGCT 2627

Qy  1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCGTGCTTTAAAACTGGCAT 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2628 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCGTGCTTTAAAACTGGCAT 2687

Qy  1981 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACACCTTTCAAGCAAGTGGC 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2688 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACACCTTTCAAGCAAGTGGC 2747

Qy  2041 AAGTTCAACAGGATTTTGTGACCAGCCGAGGCCTGGGCCTCCTTCTGCATGATTCTATCCA 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2748 AAGTTCAACAGGATTTTGTGACCAGCCGAGGCCTGGGCCTCCTTCTGCATGATTCTATCCA 2807

Qy  2101 A 2101

Db  2808 A 2808
```

ORIGIN

```
Qy     1 GAGCTATGCCTACACACAGCGTGCTTATGTCACCACCTCTGACCCTCGACCGGAGCCCATT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1095 GAGCTATGCCTACACACAGCGTGCTTATGTCACCACCTCTGACCCTCGACCGGAGCCCATT 1154

Qy    61 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1155 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1214

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1215 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1274

Qy   181 TGCTGAGGACACAATGTCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1275 TGCTGAGGACACAATGTCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1334

Qy   241 CCAGTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1335 CCAGTTCATACTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1394

Qy   301 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1395 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1454

Qy   361 AACTGAAGTACAAGAGCAGATGAATCCTCAAATTCAAGATGGGAATGGGATGCCTCAGGGTAGC 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1455 AACTGAAGTACAAGAGCAGATGAATCCTCAAATTCAAGATGGGAATGGGATGCCTCAGGGTAGC 1514

Qy   421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1515 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1574

Qy   481 GAAAGAGTTGAATGACCTGATCTTGAAGAACTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1575 GAAAGAGTTGAATGACCTGATCTTGAAGAACTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1634

Qy   541 GCCTCTTGGACCTGACCTGATCTTGAAGAACTAAAACGCCAAGTACAACACATAAGGTGCTTCA 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1635 GCCTCTTGGACCTGACCTGATCTTGAAGAACTAAAACGCCAAGTACAACACATAAGGTGCTTCA 1694

Qy   601 AGAAGATCTAGAACAAGAACAAGTCAGGGTGCAATTCTCCACTCACATGGTGGTGGTAGT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1695 AGAAGATCTAGAACAAGAACAAGTCAGGGTGCAATTCTCCACTCACATGGTGGTGGTAGT 1754

Qy   661 TGATGAATCTAGTGAGATCGGAGATCACGCAACTGCTTGCTCTTGGAAGAACAACTAAGGTATTGGG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1755 TGATGAATCTAGTGAGATCGGAGATCACGCAACTGCTTGCTCTTGGAAGAACAACTAAGGTATTGGG 1814

Qy   721 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1815 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1874

Qy   781 TTCTGACCAGTGGAAGGCTGTGCACCTTTCTGCGAAGAACAGTCGTGTGTGGCTACAGCT 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1875 CCTTCTCAAATGGCAAATGGCTCTTACTGGAAGAACAGTCGCCTTTTTAGTGCATGGCTTTCAGA 1934

Qy   841 GAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAA 900
         |||||| ||||||
Db  1935 AAAAGAAGATGCAGT----------------------------------------- 1949

Qy   901 GCAGAACCGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCAT 960
         --------------------------------------------------------
Db  1950

Qy   961 GAGTACTCTTGAGACTGTACGGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAA 1020
         ||||||||||||||  |||
Db  1950 ------------------GAACAAGAATTCACACAACTGGCTTTAAAGATCAAAATGAA 1989
```

Mon Apr 5 15:45:50 2004   us-09-845-416-12_copy_900_3000.rge

Qy   2101 A 2101
Db   2808 A 2808

Search completed: April 5, 2004, 02:40:08
Job time : 5490.46 secs

```
y  1021 ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCT 1080
            ||||||||||||||||||
b  1990 ATGTTATCAAG----------------------------------------------- 2000

y  1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACTCCGC 1140
b  2001 --------------------------------------------------------- 2000

y  1141 TGACTGGCAGAGAAAATAGATGAGACCCCTGAAAGACTTCCAGGAACTTCAAGAGGCCAC 1200
b  2001 --------------------------------------------------------- 2000

y  1201 GGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260
b  2001 --------------------------------------------------------- 2000

y  1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 1320
            |||||||||||||||||||||||||||||
b  2001 ----------TCTCGAGAAAGTCAAGGCACTTCGAGG 2027

y  1321 AGAAATTGCGCCTCTGAAAGAGAACCTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2028 AGAAATTGCGCCTCTGAAAGAGAACCTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 2087

y  1381 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 1440
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2088 CACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAG 2147

y  1441 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGCGAGCTGCATGAAGCCCACAG 1500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2148 ATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGCGAGCTGCATGAAGCCCACAG 2207

y  1501 GGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCGTGTCCAGGGTCCCTGGGAGAG 1560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2208 GGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCGTGTCCAGGGTCCCTGGGAGAG 2267

y  1561 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2268 AGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 2327

y  1621 GGACCATCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2328 GGACCATCCAAAATGACAGAGCTCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATT 2387

y  1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTGCTTGGATCT 1740
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2388 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTGCTTGGATCT 2447

y  1741 CTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCCGCCACAACCTCAAGCAAAATGACCA 1800
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2448 CTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCCGCCACAACCTCAAGCAAAATGACCA 2507

y  1801 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTATGACCGCCTGGAGCA 1860
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2508 GCCCATGGATATCCTGCAGATTATTAATTGTTTGACCACTATTATGACCGCCTGGAGCA 2567

y  1861 AGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAACTCGGCTGCT 1920
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2568 AGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGATATGTGTCTGAACTCGGCTGCT 2627

y  1921 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCGTGTCTTTTTAAAACTGGCAT 1980
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2628 GAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCGTGTCTTTTTAAAACTGGCAT 2687

y  1981 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2688 CATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2747

y  2041 AAGTTCAACAGGATTTTGTGACCAGCGGCAGCCGCAGGGCGCCCTTCCTTCTGCATGATTCATCCA 2100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b  2748 AAGTTCAACAGGATTTTGTGACCAGCGGCAGCCGCAGGGCGCCCTTCCTTCTGCATGATTCATCCA 2807
```

```
                Copyright  GenCore version 5.1.6
                Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on:   April  4, 2004, 16:52:43 ; Search time 2274.66 Seconds
                                     (without alignments)
                                     17079.796 Million cell updates/sec

Title:          US-09-845-416-10_COPY_1800_3100
Perfect score:  1301
Sequence:       1 cgactttccagcagttcaga.........ccggagctgctccaattg 1301

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27513289 seqs, 1493109276 residues

Total number of hits satisfying chosen parameters:   55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      EST:*
            1:  em_estba:*
            2:  em_esthum:*
            3:  em_estin:*
            4:  em_estmu:*
            5:  em_estov:*
            6:  em_estpl:*
            7:  em_estro:*
            8:  em_htc:*
            9:  gb_est1:*
           10:  gb_est2:*
           11:  gb_htc:*
           12:  gb_est3:*
           13:  gb_est4:*
           14:  gb_est5:*
           15:  em_estfun:*
           16:  em_estom:*
           17:  em_gss_hum:*
           18:  em_gss_inv:*
           19:  em_gss_pln:*
           20:  em_gss_vrt:*
           21:  em_gss_fun:*
           22:  em_gss_mam:*
           23:  em_gss_mus:*
           24:  em_gss_pro:*
           25:  em_gss_rod:*
           26:  em_gss_phg:*
           27:  em_gss_vrl:*
           28:  gb_gss1:*
           29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
                         SUMMARIES

          %
Result    Query
No.  Score Match Length DB Length ID          Description
----------------------------------------------------------------
  1   584  44.9   824    9  AL556247    AL556247 AL556247
  2   583  44.8   777   14  CD653550    CD653550 AGENCOURT
  3   582.4 44.8 1098   13  BX365572    BX365572 BX365572
  4   573  44.0   620   13  BQ640063    BQ640063 he23g04.y
```

```
  5  571.4 43.9  571.4 14  801   CB991394   CB991394 AGENCOURT
  6  537.8 41.3  537.8 12  770   BG719710   BG719710 60269O430
  7    535 41.1    535 11 3753   AK081426   AK081426 Mus muscu
  8  526.4 40.5  526.4 11 1298   AK087829   AK087829 Mus muscu
  9  526.4 40.5  526.4 11 2135   AK013510   AK013510 Mus muscu
 10  526.4 40.5  526.4 11 4437   AK036936   AK036936 Mus muscu
 11  524.8 40.3  524.8 11 1384   AK075809   AK075809 Mus muscu
 12    511 39.3    511 14  797   CB960722   CB960722 AGENCOURT
 13  508.2 39.1  508.2 29 2874   AY408546   AY408546 Homo sapi
 14  502.8 38.6  502.8 10  652   BB629984   BB629984 BB629984
 15    497 38.2    497 29 2874   AY408548   AY408548 Mus muscu
 16  491.6 37.8  491.6 10  704   BB610411   BB610411 BB610411
 17  489.8 37.6  489.8 14  508   CD701871   CD701871 EST18395
 18  479.6 36.9  479.6  9  854   AI196693   AI196693 ui53e10.y
 19  470.0 36.2  470.0 14  717   CB527785   CB527785 UI-M-FY0-
 20  461.2 35.4  461.2 29  542   CG594848   CG594848 OST253426
 21  457.2 35.1  457.2 29  513   CG605968   CG605968 OST283399
 22  455.8 35.0  455.8 13  591   BX490860   BX490860 DKFZp686F
 23  453.8 34.9  453.8 29  508   CG606947   CG606947 OST285335
 24    439 33.7    439 13  578   BX485574   BX485574 DKFZp586L
 25  436.6 33.6  436.6 13  665   BY742604   BY742604 BY742604
 26  431.4 33.2  431.4 10  898   BF182065   BF182065 60180460́4
 27  427.8 32.9  427.8 14  843   CA988247   CA988247 AGENCOURT
 28  426.6 32.8  426.6 11 3051   BC036095   BC036095 Homo sapi
 29  422.6 32.5  422.6 29  520   CG606443   CG606443 OST284316
 30  420.2 32.3  420.2 14  512   CF162938   CF162938 B0718G07-
 31  419.4 32.2  419.4 29  488   CG597965   CG597965 OST261937
 32    416 32.0    416 29  493   CG604818   CG604818 OST280872
 33  410.4 31.5  410.4 10  599   BB666688   BB666688 BB666688
 34  407.2 31.3  407.2 14  554   CB613696   CB613696 AMGNNUC:N
 35  405.4 31.2  405.4 12  662   BJ626491   BJ626491 BJ626491
 36  404.4 31.1  404.4 14  493   CA888041   CA888041 B0142C06-
 37  397.4 30.5  397.4 14  495   CA894775   CA894775 B0187G06-
 38  392.2 30.1  392.2 13  650   BY714491   BY714491 BY714491
 39    389 29.6    389 13  794   BU424348   BU424348 603234189
 40  385.4 29.6  385.4  9  677   AL641565   AL641565 AL641565
 41  383.4 29.5  383.4 14  483   CA893902   CA893902 B0182B01-
 42  381.8 29.3  381.8  9  423   AA460476   AA460476 zx6le10.r
 43  380.4 29.2  380.4  9  622   AL796733   AL796733 AL796733
 44  375.4 28.9  375.4  9  502   AL871560   AL871560 AL871560
 45  365.4 28.1  365.4 14  655   CA558919   CA558919 K025lG03-
```

```
                         ALIGNMENTS

RESULT 1
AL556247
LOCUS       AL556247                824 bp    mRNA    linear    EST 31-MAY-2003
DEFINITION  AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CSODK001YB17 5-PRIME, mRNA sequence.
ACCESSION   AL556247
VERSION     AL556247.2  GI:31278051
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 824)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12898746.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1955.r
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DK001CA09QPl.
FEATURES             Location/Qualifiers
```

```
     source          1..824
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK001YB17"
                     /cell_type="HELA"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match        44.9%;  Score 584;  DB 9;  Length 824;
Best Local Similarity   100.0%;  Pred. No. 6.1e-163;
Matches 584;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    718 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTT 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTT 254

Qy    778 AGCTGACCTGAATAATGTCAGATTCTCAGCTGACTGCTATAGGACTGCCATGAAACTCCGAAGACT 837
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 AGCTGACCTGAATAATGTCAGATTCTCAGCTGACTGCTATAGGACTGCCATGAAACTCCGAAGACT 314

Qy    838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCAGTGTGATGCCTTGGACCA 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCAGTGTGATGCCTTGGACCA 374

Qy    898 GCACAACCTCAAGCAAAATGACCAGCCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 957
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 GCACAACCTCAAGCAAAATGACCAGCCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGAC 434

Qy    958 CACTATTATGACCGGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCCTCTGCGT 1017
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 CACTATTATGACCGGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCCTCTGCGT 494

Qy   1018 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 1077
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 554

Qy   1078 TGTCCTGTCTTTAAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAACAGTA 1137
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 TGTCCTGTCTTTAAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAACAGTA 614

Qy   1138 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCCAGGCTGGG 1197
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCCAGGCTGGG 674

Qy   1198 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGG 1257
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGG 734

Qy   1258 GGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          ||||||||||||||||||||||||||||||||||||||||||||
Db    735 GGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 778
```

```
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Daniela S. Gerhard, Ph.D.
         Office of Cancer Genomics
         National Cancer Institute / NIH
         Bldg. 31 Rm10A07 Bethesda, MD 20892
         Email: cgapbs-r@mail.nih.gov
         Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
         cDNA Library Preparation: Yulan Piao and Minoru Ko
         cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Agencourt Bioscience Corporation
         Clone distribution: MGC clone distribution information
         can be found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: NDAM513 row: a  column: 23
         High quality sequence stop: 715.

FEATURES             Location/Qualifiers
     source          1..777
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30426742"
                     /tissue_type="Embryonic Stem cells"
                     /cell_line="WA01"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library
                     (Long)"
                     /note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
                     This is a long-transcript enriched cDNA library (Genome
                     Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
                     cell line . Undifferentiated human ES cell line WA01/H1
                     was obtained from WiCell Research Institute, Inc.
                     Madison, Wi. cultured according to their instructions, on
                     MEF feeders. They formed round colonies with defined edges
                     and were positive for alkaline phosphatase, SSEA-4, OCT3,
                     OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45.  They are
                     negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
                     TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days
                     after plating), the ES cells from 4 X 6cm dishes were
                     treated with 1 mg/ml collagenase, type IV
                     (Invitrogen/GIBCO) for 5-10 min and gently scraped off
                     with 5 ml pipette.  RNA was purified with TRIzol Reagent
                     from Invitrogen Protocol ref: Genome Res. 11: 1553-1558
                     (2001). [PMID:11544199]) Double-stranded cDNAs were
                     synthesized with an oligo(dT) primer [Invitrogen:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3'] from
                     3.4g of total RNA, treated with T4 DNA polymerase, and
                     purified by ethanol-precipitation. The cDNAs were ligated
                     to Lone-linker LL-Sal4, purified by phenol/chloroform
                     extraction, and separated from free linkers by
                     Centricon-100 column. Then, the cDNAs were amplified by
                     long-range high fidelity PCR using Ex Taq polymerase
                     (Takara) with a primer Sal4-S for 25 cycles. The products
                     were purified by phenol/chloroform extraction and
                     Centricon-100 column. The cDNAs were digested with SalI
                     and NotI enzymes and cloned into SalI/NotI site of
                     pCMV-SPORT6 plasmid vector. The average insert size is
                     about 3.6kb."

ORIGIN

Query Match         44.8%;  Score 583;  DB 14;  Length 777;
Best Local Similarity   99.8%;  Pred. No. 1.2e-162;
Matches 583;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy    718 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTT 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTT 196

Qy    778 AGCTGACCTGAATAATGTCAGATTCTCGCTGACTGCCATTATAGGACTGCCATGAAACTCCGAAGACT 837
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    197 AGCTGACCTGAATAATGTCAGATTCTCAGCTGACTGCCATTATAGGACTGCCATGAAACTCCGAAGACT 256

Qy    838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCAGTGTGATGCCTTGGACCA 897
```

```
RESULT 2
CD653550

LOCUS       CD653550                 777 bp    mRNA    linear   EST 18-JUN-2003
DEFINITION  AGENCOURT_14553000 NIA Human H1 Embryonic Stem Cell cDNA Library
            (Long) Homo sapiens cDNA clone IMAGE:30426742 5', mRNA sequence.
ACCESSION   CD653550
VERSION     CD653550.1  GI:31892067
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 777)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
```

Best Local Similarity 99.8%;   Pred. No. 2.3e-162;
Matches 583;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

```
Qy   718 CCACGAGACTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT 777
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   204 CCACGAGACTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT 263

Qy   778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGAGACTGCCATGAAACTCCGAAGACT 837
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   264 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGAGACTGCCATGAAACTCCGAAGACT 323

Qy   838 GCGAAGAGGCCCTTGCTTGGATCCTTGAGCCGTCAGCTGCATGTGATGCCTTGGACCA 897
         |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   324 GCGAAGAGGCCCTTGCTTGGKTCTTGAGCCGTCAGCTGCATGTGATGCCTTGGACCA 383

Qy   898 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCGCAGATATCCTGCAGATTAATGTTTGAC 957
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   384 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCGCAGATATCCTGCAGATTAATGTTTGAC 443

Qy   958 CACTATTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTGCGT 1017
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   444 CACTATTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTGCGT 503

Qy  1018 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACGAGGAGGATCCG 1077
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   504 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACGAGGAGGATCCG 563

Qy  1078 TGTCCGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGAACAAGTA 1137
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   564 TGTCCGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGAACAAGTA 623

Qy  1138 CAGATACCTTTTCTGCAAGCAAGTGGCAAGTGGCAAGTTCAAAATTCCAAGACAGTGGGTGCATCCTTTGG 1197
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   624 CAGATACCTTTTCTGCAAGCAAGTGGCAAGTGGCAAGTTCAAAATTCCAAGACAGTGGGTGCATCCTTTGG 683

Qy  1198 CCTCCTTCTGCATGATTCATTCATCCAAATTCCAAGACAGTGTCCGGAGCTGCTTCCAATTTG 1257
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   684 CCTCCTTCTGCATGATTCATTCATCCAAATTCCAAGACAGTGTCCGGAGCTGCTTCCAATTTG 743

Qy  1258 GGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
         ||||||||||||||||||||||||||||||||||||||||||||
Db   744 GGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 787
```

RESULT 4
BQ640063
LOCUS   BQ640063   620 bp   mRNA   linear   EST 15-JUL-2002
DEFINITION he23g04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
            Homo sapiens cDNA clone he23g04 5', mRNA sequence.
ACCESSION   BQ640063
VERSION   BQ640063.1 GI:21764522
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 620)
  AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE   Expressed sequence tag analysis of human retina for the NEIBank
            Project: Retbindin, an abundant, novel retinal cDNA and alternative
            splicing of other retina-preferred gene transcripts
  JOURNAL   Mol. Vis. 8 (4), 196-204 (2002)
  MEDLINE   22103461
   PUBMED   12107411
COMMENT   Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 23 row: g column: 04

```
Db   257                   |||||||||||||||||||||||||||||||||||||||||||
         GCAGAAGAGGCCCTTGCTTGGATCCTTGAGCCGTCAGCTGCTGCAGTGATGATGCATGTGGACCA 316

Qy   898 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCGCAGATATCCTGCAGATTAATGTTTGAC 957
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   317 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCGCAGATATCCTGCAGATTAATGTTTGAC 376

Qy   958 CACTATTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTGCGT 1017
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   377 CACTATTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTGCGT 436

Qy  1018 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACGAGGAGGATCCG 1077
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACGAGGAGGATCCG 496

Qy  1078 TGTCCGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGAACAAGTA 1137
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   497 TGTCCGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGAACAAGTA 556

Qy  1138 CAGATACCTTTTCTGCAAGCAAGTGGCAAGTGGCAAGTTCAAAATTCCAAGACAGTGGGTGCATCCTTTGG 1197
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   557 CAGATACCTTTTCTGCAAGCAAGTGGCAAGTGGCAAGTTCAAAATTCCAAGACAGTGGGTGCATCCTTTGG 616

Qy  1198 CCTCCTTCTGCATGATTCATTCATCCAAATTCCAAGACAGTGTCCGGAGCTGCTTCCAATTTG 1257
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   617 CCTCCTTCTGCATGATTCATTCATCCAAATTCCAAGACAGTGTCCGGAGCTGCTTCCAATTTG 676

Qy  1258 GGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
         ||  |||||||||||||||||||||||||||||||||||||||
Db   677 GGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCNATTTG 720
```

RESULT 3
BX365572
LOCUS   BX365572   1098 bp   mRNA   linear   EST 05-MAY-2003
DEFINITION BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK001YB21 5-PRIME, mRNA sequence.
ACCESSION   BX365572
VERSION   BX365572.1 GI:30366927
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1098)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1955.r
            Contact : Feng Liang Email : fliang@lifetech.com URL ;
            http://fulllength.invitrogen.com/ inVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS1AK001ZB11QP1.
FEATURES   Location/Qualifiers
     source   1..1098
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK001YB21"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with NotI and cloned into the NotI and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 44.8%;   Score 582.4;   DB 13;   Length 1098;

RESULT 5

```
LOCUS       CB991394                801 bp    mRNA      linear   EST 01-MAY-2003
DEFINITION  AGENCOURT_13627932 NIH_MGC_148 Homo sapiens cDNA clone
            IMAGE:30336570 5', mRNA sequence.
ACCESSION   CB991394
VERSION     CB991394.1  GI:30285818
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM360  row: d  column: 19
            High quality sequence stop: 621.
FEATURES             Location/Qualifiers
     source          1..801
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3033570"
                     /tissue_type="pre-eclamptic placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_148"
                     /note="Organ: placenta; Vector: pBluescriptR; Site 1:
                     all-XhoI; Site 2: BamH; Library is oligo-dT primed and
                     directionally cloned using primer
                     5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
                     size 2.3 kb  and normalized to ROT 5. This is a primary
                     library enriched for full-lenght clones and constructed
                     using the Cap-trapper method (Carninci, in preparation).
                     Library constructed by M. Brownstein (NIMH/NHGRI,
                     National Institutes of Health). Note: this is a NIH_MGC
                     Library."
ORIGIN
```

Query Match   43.9%;   Score 571.4;   DB 14;   Length 801;
Best Local Similarity 99.7%;   Pred. No. 3.5e-159;
Matches   583; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

```
Qy     718 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTT 777
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTT 199

Qy     778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 259

Qy     838 GCAGAGGCCCTTGCTTGGATCCTGTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA 897
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 GCAGAGGCCCTTGCTTGGATCCTGTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA 319

Qy     898 GCACAACCTCAAGCAAAATGACACCAGCCCATGGGATATCCTGCAGATTATTAATTGTTGAC 957
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 GCACAACCTCAAGCAAAATGACACCAGCCCATGGGATATCCTGCAGATTATTAATTGTTGAC 379

Qy     958 CACTATTTATGACCGCCTGGAGCAAGAGACAAGAGCACAACAATTTGGTCAACGTCCCTCTCGCGT 1017
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     380 CACTATTTATGACCGCCTGGAGCAAGAGACAAGAGCACAACAATTTGGTCAACGTCCCTCTCGCGT 439
```

Seq primer: M13RP1 reverse primer (ABI).

```
FEATURES             Location/Qualifiers
     source          1..620
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="he23g04"
                     /tissue_type="Retina"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Retina cDNA (Un-normalized,
                     unamplified): hd/he"
                     /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
                     was dissected from two 80 year old donors with no observed
                     eye disease. 100ug of total RNA was used for library
                     construction. A directionally cloned cDNA library in the
                     pSPORT1 vector (Life Technologies) was constructed at the
                     Bioserve Biotechnology (Laurel MD) essentially following
                     the protocols of the SuperScript Plasmid System full
                     details of which are contained in the manufacture's
                     instruction manual (http://www.lifetech.com/). First
                     strand synthesis was carried out using a Not I
                     primer-adapter
                     [5'-pGACTAGTTCTAGATCGGCGAGCGCGCGCCC(T)15-3']. EST analysis
                     was performed on the unamplified library at the NIH
                     Intramural Sequencing Center (NISC)."
ORIGIN
```

Query Match   44.0%;   Score 573;   DB 13;   Length 620;
Best Local Similarity 100.0%;   Pred. No. 9.7e-160;
Matches   573; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     729 AAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGTGACCTGA 788
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGTGACCTGA 60

Qy     789 ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCC 848
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCC 120

Qy     849 TTTGCTTGGATCCTCTGAGCCTGTCAGCTGCTGCATGTGATGCCTTGGACAGCACAACCTCA 908
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTTGCTTGGATCCTCTGAGCCTGTCAGCTGCTGCATGTGATGCCTTGGACAGCACAACCTCA 180

Qy     909 AGCAAAAATGACCAGCCCATGGATATCCTGCAGATATCCTGAGATTATTAATTGTTGACCACTATTTATG 968
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGCAAAAATGACCAGCCCATGGATATCCTGCAGATATCCTGAGATTATTAATTGTTGACCACTATTTATG 240

Qy     969 ACCGGCCTGGAGCAAGAGACAAGAGCACAACAATTTGGTCAACGTCCCTCTCGCTGGATATGTGTC 1028
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACCGGCCTGGAGCAAGAGACAAGAGCACAACAATTTGGTCAACGTCCCTCTCGCTGGATATGTGTC 300

Qy    1029 TGAACTGGCTGCTGAATGTTTATGATACCGGACGAACAGGGAGGATCCGTGTCCTGTCTT 1088
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACTGGCTGCTGAATGTTTATGATACCGGACGAACAGGGAGGATCCGTGTCCTGTCTT 360

Qy    1089 TTAAAACTGGACATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGCAACAGATACCTTT 1148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTAAAACTGGACATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGCAACAGATACCTTT 420

Qy    1149 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGC 1208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCTCCTTCTGC 480

Qy    1209 ATGATTCTATCCAAATTCCAAGAACAGAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACA 1268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ATGATTCTATCCAAATTCCAAGAACAGAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACA 540

Qy    1269 TTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
           |||||||||||||||||||||||||||||||||
Db     541 TTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 573
```

```
Qy    718  CCACGAGACTCAAACAACTTGCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTT  777
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    150  CCACGAGACTCAAACAACTTGCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTT  209

Qy    778  AGCTGACCTGAATAATGTCAGATTCCTCAGTTATAGGACACATTTGGAAGACAAGTA  837
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    210  AGCTGACCTGAATAATGTCAGATTCCTCAGTTATAGGACACATTTGGAAGACAAGTA  269

Qy    838  GCAGAAGGCCCTTGCTTGGATCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA  897
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270  GCAGAAGGCCCTTGCTTGGATCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA  329

Qy    898  GCACAA--CCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTG  955
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330  GCACAATCCTCAAGTCAGAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTG  389

Qy    956  ACCACTATTTATGACCGGCTGAGCCTGAGCAAGACAACAATTTGGTCAACGTCCCTCTCGC  1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390  ACCACTATTTATGACCGGCTGAGCCTGAGCAAGACAACAATTTGGTCAACGTCCCTCTCGC  449

Qy    1016 GTGGATATGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATC  1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450  GTGGATATGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATC  509

Qy    1076 CGTGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAG  1135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510  CGTGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAG  569

Qy    1136 TACAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTG  1195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570  TACAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGA-TTTGTGACCAGCGCAGGCTG  628

Qy    1196 GGCCTCCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTT  1255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    629  GGCCTCCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTT  688

Qy    1256 GGGGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAAT  1298
           ||||||||||||||||||||||||||||||||||||||||
Db    689  GGGGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATT  731
```

```
RESULT 7
AK081426
LOCUS       AK081426        3753 bp    mRNA    linear    HTC 20-SEP-2003
DEFINITION  Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
            library, clone:C130016K19 product:dystrophin related protein 2,
            full insert sequence.
ACCESSION   AK081426
VERSION     AK081426.1    GI:26349154
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
  PUBMED    10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
  PUBMED    11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
```

```
                                            /           1018 GGATATGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG  1077
                                            >                440 GGATATGTGTCGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG  499

                                            /           1078 TGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTA  1137
                                            >                500 TGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTA  559

                                            /           1138 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG  1197
                                            >                560 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGG  619

                                            /           1198 CCTCCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTG  1257
                                            >                620 CCTCCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTG  679

                                            /           1258 GGGCAGTAACATTGAGCCAAGTGTCC-GGAGCTGCTTCCAATTTG  1301
                                            >                680 GGGCAGTAACATTGAGCCAAGTGTCCGGGAGCTGCTTCCAATTTG  724
```

```
RESULT 6
BG719710
LOCUS       BG719710   602690430F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822807 5',
            mRNA sequence.                             770 bp    mRNA    linear    EST 08-MAY-2001
DEFINITION  
ACCESSION   BG719710
VERSION     BG719710.1    GI:13998897
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
            National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE     Unpublished (1999)
  JOURNAL   
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10731 row: 1 column: 08
            High quality sequence stop: 767.
FEATURES             Location/Qualifiers
     source          1..770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4822807"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTVN-3', size-selected for average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation). Library constructed by M. Brownstein
                     (NIH/NHGRI), National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

Query Match      41.3%;   Score 537.8;   DB 12;   Length 770;
Best Local Similarity   97.4%;   Pred. No. 4e-149;
Matches  568;   Conservative   0;   Mismatches  12;   Indels  3;   Gaps  2;
```

```
          Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
          Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,K., Matsura,S., Kawai,J.,
          Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system-384-format
          sequencing pipeline with 384 multicapillary sequencer
  JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE 20530913
   PUBMED 11076861
REFERENCE 4
  AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
          FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
  AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
  JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6  (bases 1 to 3753)
  AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
          Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
          Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
          Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
          Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
          Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
          Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
          Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
          Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
          Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
          Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
          Physical and Chemical Research (RIKEN), Laboratory for Genome
          Exploration Research Group, RIKEN Genomic Sciences Center (GSC);
          RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
          Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
          URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
          Fax:81-45-503-9216)
  COMMENT cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site for further details.
          URL:http://genome.gsc.riken.go.jp/
          URL:http://fantom.gsc.riken.go.jp/.
 FEATURES       Location/Qualifiers
     source     1..3753
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="FANTOM_DB:C130016K19"
                /db_xref="MGI:2413800"
                /db_xref="taxon:10090"
                /clone="C130016K19"
                /tissue_type="head"
                /clone_lib="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="16 days embryo"
     CDS        181..3054
                /note="unnamed protein product; dystrophin related protein
                2 (MGD|MGI:107432, GB|U43520, evidence: BLASTN, 99%,
                match=807)
                putative"
                /codon_start=1
                /protein_id="BAC38217.1"
                /db_xref="GI:26349155"
                /translation="MQPLVMGGCPYTLPRCHEWHAADRFHHSSSLRNTCPQPQVRAAV
                TIPAPPWDGAGDPCLSPKLLNGTVGATGPLEPSAMNLCWNEIKKKSHNLRARLEAFSD
                LSGKLQLPLREIIDWLSQXDEBLSAQLPLQGDVALVQQEKETHAAFMEEVKSKGPYIS
                SVLESAQFLSQHPFEELEESQSESKDTSPRQRIQNLSRFVWKQATVASELWEKLTAR
                CVDQHRHIEHTLEHLLEIQGAMEELSSTITQAEGVRATWEPIGDLFIDSLPEHIQAIK
                LFKEEFSPVKDGVGVKLVNDLAHQLAISDVHLSMENSRALEQININRWKQLQVSVAERLKQ
                LQDAHRDFPGPSGDFHFLSTSVQVPWERAFSPNKVPYYINHQAQTTCWDHPKMWTELYQTL
                ADLNNIKFSAYRTAMKLRRVQKALRLDJVTLTLTLALEIFNEHDLQASEHVMDVVEVIHC
                LTALYERLEBERGILVNVPLCVDMSLNWLLNVFDSGRSGKMRALSFKTGIACLCGTEV
                KEKLQYLFSQVANSGSQCDRHLGALLHEAIQVPRQLEVAAFGGSNVEPSVRSCRF
                STGKPVIEASQFLEWVNLEPQSMVVLAVLHRVTIAEQVKHQTKCSICRQCPIKGFRYR
                SLKQFNVDICQTCFLTGRASKGNKLHYPIMEYYTPTTSSENMRDFATTLKNKFRSKQY
                FSKHPQRGYLPVQSVLESDCSETPASSPMLPHADTHSRIEHFASRLAEMESQNCSFPN
                DSLSPDDSIDEDQYLLRHSSPITDREPAFGQQAPCSMATESKGELEKILAHLEDENRI
                LQGELRIRLKWQHEAAAEAPTLVEGSARATPDHRNEBLLAEARILRQHSKSRLETRMQIL
                EDHNKQLESQLQRIRELLLQPPSESDGNGSAGSSLASSPRQSEGSHPREKGQTTPDTE
                VADDVGSKSQDVSLCLEDIMEKLRHAFPSVRSSDVTANTLLAS"
     polyA_signal  3731..3736
                /note="putative"
     polyA_site    3753
                /note="putative"
 ORIGIN

 Query Match      1..3736          Score 535;  DB 11;   Length 3753;
 Best Local Similarity   41.1%;   Pred. No. 8.8e-148;
 Matches  817;  Conservative   0;  Mismatches 470;  Indels   0;  Gaps   0;

 Qy   14 GTTCAGAAGCAGCAAGCGATGTACATAGGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCT   73
            ||||||||||  |||||||   |||| ||  |  ||||||  ||||| || ||| ||
 Db  592 GTACACAGGAGGAGGACACACATGCAGCCCTTCATGGAAGAAGTCAAGTCTAAGGGCCCC  651

 Qy   74 GTAATCATGAGTACTCCTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGA  133
            ||| | | || || |   ||   |||  || | |||||||   ||  |||   ||||
 Db  652 TACATCCTCTGTGCTTGAATCCGCCCAGGCTTTTCTGTCCCAGCATCCATTTGAAGAAGA  711

 Qy  134 CTAGAGAAACTCACCAGGAGCCCAGAGAGTGCTCCTGGAGAGAGCGCGGATTCAAAACCTT  193
            ||||||||| ||  ||||| |||  || | || |  ||||   | |||  ||||  ||
 Db  712 TTAGAGGAGTCTCAATCTGAGAGTAAAGATACCTCCCCCAGAGAATGAGTGGGAAAATTGA  771

 Qy  194 ACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAATTGAACCTG  253
            || ||||  || | || | || ||   |||  | || ||| |   || || | ||
 Db  772 AGTCGGCTTTGTATGGAAGCAGGCAACAGTGCCAACAGTGCCGTGAGCTGTGGGAAGCTGACAGCC  831

 Qy  254 CACTCCCGACTGGCAGAGAAAAATAGATGAGATGCCACATTGAGCATATCTGGAACTTCAA  313
            |   |  ||||  ||   | |   ||| |  ||||||   ||  |||  |||||| | |||
 Db  832 CGCTGTGTAGATCAGCATCGGCCACATTGAGCATCTATTGGAGATCCAA  891

 Qy  314 GAGGCCACGGATGAGCTGGACCTCAAGTCGCCGCAAGCTGAGGTGATCAAGGATCCTGG  373
            |||||||||||| |||||||  | || ||| ||| ||||| ||||||||||| |||||
 Db  892 GGGGCAATGAGGAGGAATTGAGCAGTACTTTGACCCAGCAGCAGGGAGTCCGAGCCACGTGG  951

 Qy  374 CAGCCCGTGGGCGATCCTCCATTGACTCCTCCAAGATCACCTCGAGAAAGTCAAGGCCA  433
            ||||||  |||| |||| |   || || ||||||| |   |||| || ||||  |||
 Db  952 GAGCCCATAGGAGAGATCTTTTATTGATTCCCTGCCAGAGCATATCCAAGCCATCAAGCTA  1011

 Qy  434 CTTCGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGCTCGC  493
            ||  || |||| |||| || ||||  ||||| |||  |||| | ||| |||| |||
 Db 1012 TTCAAAGAAGAAGAATTCTCTCCTGTGAAAGATGGGGTGAAGTTAGTGAATGATCTGGCCCAC  1071

 Qy  494 CAGCTTACCACTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTTCAGGAGACCTG  553
            ||||| |||  |||||| | |  | ||| ||   || ||||| ||||||||||   |
 Db 1072 CAGCTTGCCATCTCTGATGTGCACTTGTCAATGGAGAATTCCAGGGCTCTGGAACAGATC  1131

 Qy  554 AACAACCAGATGAAGCTTCTGCAGGTGGCCGTGGCCTCGAGGAGCCAGTCGCAGCTGCATGAA  613
            |||  ||||| ||||| || || |||| || || ||  || | || || ||| |||
 Db 1132 ACATCCCGTGAAAGCACAGCTCCAGGTGTCAGTGGCTGAGAGGCTTAAGCAACTCAACTCCAGGAT  1191

 Qy  614 GCCCACAGGGACTTGGTCCAGCATCTCAGCACTTCTTCTCCCTGCCCACGTCTGCCCAGGGTCCC  673
            ||||| | ||| | |||  |||| |||||||| | | || ||   ||| || ||  | |||
 Db 1192 GCCCACCGGGGACTTGGGCCTTGGGCACAGCACTTCCTCTCCACTTCGTCCAAGTTCCC  1251

 Qy  674 TGGGAGGAGAGCCATCTGCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA  733
            |||| ||||||| ||| |    || | ||||| ||||| | |||| || ||||| |||||
 Db 1252 TGGGAAAGAGAGCATTTTCTCCCAATAAAGTCCCCTACTACAGTCCTTAGCTGACCTGAATAAT  1311

 Qy  734 ACTTGCTGGGACCATCCGAGTCAGTGAGTTAGCTGACAGACCCTAGCACCCTAGCCATGAACC  793
            || | ||||||| ||| || | |||| |  | |   ||   || | || ||  |||| ||
 Db 1312 ACATGCTGGGACCATCCTCAAGATGACTGAGTTATCCCAAAACCCTAGCCGATCTGAACAAC  1371
```

TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED   11076861
REFERENCE   4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1298)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,H., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT   cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES   Location/Qualifiers
source   1..1298
     /organism="Mus musculus"
     /mol_type="mRNA"
     /strain="C57BL/6J"
     /db_xref="FANTOM_DB:E330026B12"
     /db_xref="MGI:2426619"
     /db_xref="taxon:10090"
     /clone="E330026B12"
     /sex="female"
     /tissue_type="ovary"
     /clone_lib="RIKEN full-length enriched mouse cDNA library"
     /dev_stage="2 days pregnant adult"
CDS   120..1142
     /note="unnamed protein product; dystrophin, muscular
dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN,
100%, match=999)
putative"
     /codon_start=1
     /protein_id="BAC40019.1"
     /db_xref="GI:26352778"
     /translation="MREHLKGHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKL
RRLQKALCLDLLSLSAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLREQERNNLVN
VPLCVDMCLNWLLNVYDTGTRIRVLSFKTGIISLCKAHLEDKYRYLPKQVASSTGF
CDQRRIGLLLHDSIQIPROLGEVASFGGSNIEPSVRSCFQFANNKPEIEAALPLDWMR
LEFQSMWLPVLHRVAAAETAKHQAKCNICKECPIIGFRYRSLGHFNYDICQSCFFSG
RVAKGHKMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVOTVLE
GDNMET"

RESULT 8
AK087829
LOCUS   AK087829   1298 bp   mRNA   linear   HTC 20-SEP-2003
DEFINITION   Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
full-length enriched library, clone:E330026B12 product:dystrophin,
muscular dystrophy full-length cDNA, full insert sequence.
ACCESSION   AK087829
VERSION   AK087829.1   GI:26352777
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE   High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED   10349636
REFERENCE   2
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED   11042159
REFERENCE   3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

y   794 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCGAAGGCCCTTGC 853
     |||||||| | ||||||||||||| |||| ||||||||| || || | | |||||||
b   1372 ATTAAGTTCTCAGCTTACCGCACTGCCATGAAGCTCCGAGAGTCCAGAAGGCCCTGCGC 1431

y   854 TTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAA 913
     |||| | |||||||||||||||||||| ||||| |||||| |||||||||||||||||
b   1432 CTGGGATCTGGTAACTTAACTACACAGCTCTGGAGATCTTCAATGACATGACTTGCAGGCC 1491

y   914 AATGACCAGCCCATGAGATATCCTGCAGATTATTAATTGTTGACCACTATTATGACCGC 973
     |||||| || || ||||| | ||||||||| ||||| ||||||||||||||||||||
b   1492 AGTGAACATGTGATGGATGTGGTGGAGGTCATTCACTGCTTGACTGCCTTGTATGAACGA 1551

y   974 CTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCCTTCGCGTGGGATATGTGTCTGAAC 1033
     |||||||| |||||| | || | ||||||||| ||| ||||||||||||| || |||||
b   1552 CTGGGAGGAGGAAAAGAGGCATCCTGGTCAATGTCGTGTAGACATGAGCCTCAAC 1611

y   1034 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGGATCCGTGTCCTGTCTTTTAAA 1093
     ||||| ||||||| |||| || | |||| ||||| ||| | ||||| | |||| |||
b   1612 TGGCTCCTCAATGTTTTTGATAGTGTCGCAGTGAAAGATGCGGAGCATTGTCCTTTAAG 1671

y   1094 ACTGGCATCATTCCCTGTGTAAAGCACACATTTGGAAGACAAGCAAGTACAGATACCTTTTCAAG 1153
     |||||| |||| ||| |||||| ||| |||| || ||| || ||||| | || ||||||
b   1672 ACTGGCATCCGCATGCCTGTGTGGCACGGAAGTGAAGTGAAAGAAAAACTTCAGTATCTCTTCAGC 1731

y   1154 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCCTCCTTCTGCATGAT 1213
     |||||| | ||||||||||||| |||||| ||| ||||| ||| | |||||| ||||
b   1732 CAAGTAGCCCAATTCAGGCAGCCAGTGTGATCAACGCCATCTCGGTGCCCTGCTTCATGAA 1791

y   1214 TCTATCCAAATTCCAAGACAGTAGCAGTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAG 1273
     |||||| |||| ||| | || | ||||||||||||| ||||||| ||||||||| |||||||
b   1792 GCCATTCAAGTGCCCGTCAGCTGGGTGAAGTGGCAGCATTTGGGGGCAGCAATGTGGAG 1851

y   1274 CCAAGTGTCCGGAGCTGCTTCCAATTT 1300
     |||||| || | || |||||||||
b   1852 CCCAGTGTCCGTCCGTAGCTGCTTTCGCTTT 1878

ORIGIN

```
Query Match        40.5%;   Score 526.4;  DB 11;  Length 1298;
Best Local Similarity  93.8%;  Pred. No. 1.5e-145;
Matches  548;  Conservative   0;  Mismatches  36;  Indels   0;  Gaps   0;

?y    718 CCACGAGACTCAAACAACTGCTGGGACCATCCCAAAATGCACAGAGCTCACCAGTCTTT 777
          ||||||||||| |||||| ||||||  |||||||| |||||||||||||||||||||||
?b    140 CCACGAGACCCAAACCACTTGTTGGGACCACCCAAAATGCACAGAGCTCTACCAGTCTTT 199

?y    778 AGCTGACCTGAATAATGCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
          |||||||| || |||||||||||| |||| |||||||||| |||||||||||||||
?b    200 AGCTGACCTGAATAATGTCAGTTCAGTTCTCCGCGTATAGGACTGCCATGAAGCTCAGAAGGCT 259

?y    838 GCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCCATGTGATGCCTTGGACCA 897
          |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
?b    260 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCCATGTGATGCCCTGGACCA 319

?y    898 GCACAACCTCAAGCAAAATGACCAGCCATGGATATCCTGCAGATTATTAATGTTTGAC 957
          |||||||||||||||||| |||||||||||||||||||||||||| |||| |||||
?b    320 GCACAACCTCAAGCAAAATGACCAGCCATGGATATCCTGCAGATAATTAACTGTTTGAC 379

?y    958 CACTATTTATGACCGCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCCTGCGT 1017
          |||| |||||||| ||||||||||||||||||| |||||||||||||||||||||||
?b    380 TACAATTTATGATCGTCTGGAGCAAGAGCACAACAATCTGTCAATGTCCCTCTCTGTGT 439

?y   1018 GGATATGTGTCGAACTGGCTGCTGAATGTTATGTTATGATACGGGACGAACAGGGAGGATCCG 1077
          |||||||||||||||||| |||| ||  |||||||||||||||||||||||||||||||
?b    440 GGATATGTGTCCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 499

?y   1078 TGTCCGTCTTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTGGAAGACAAGTA 1137
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||
?b    500 TGTCCGTCTTTTAAAACTGGCATCATTTCCTGTGTAAAGCACACTTGGAAGACAAGTA 559

?y   1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGACAGGCTGGG 1197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
?b    560 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAGCGTAGGCTCGG 619

?y   1198 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAGTTGCATCCTTTGG 1257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
?b    620 TCTTCTTCTGCATGATTCTTATTCAATTCAAATCCAAGACAGTTGGGTGAGTTGCTTCCTTTGG 679

?y   1258 GGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||||||||||||||
?b    680 GGGCAGTAGTAACATTGAGCCGGAGTGTCAGGAGCTGCTTCCAATTTG 723
```

RESULT 9
AK013510
LOCUS   AK013510   2135 bp   mRNA   linear   HTC 20-SEP-2003
DEFINITION   Mus musculus adult male hippocampus cDNA, RIKEN full-length
             enriched library, clone:2900010C03 product:dystrophin, muscular
             dystrophy, full insert sequence.
ACCESSION   AK013510
VERSION   AK013510.1 GI:12850899
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
             Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
             Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
             Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
             Yamamoto,R., Matsumoto.H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
             Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
             Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
             Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system-384-format
             sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
             FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
             Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 2135)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
             Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
             Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
             Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
             Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
             Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
             Saito,R., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
             Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
             Suzuki,H., Tagami,Y., Tagawa,A., Takahashi,F., Tanaka,T.,
             Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
             Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
             further details.
             cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues. First strand cDNA was primed with a primer
             [5' GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
             prepared by using trehalose thermo-activated reverse transcriptase
             and subsequently enriched for full-length by cap-trapper. Second
             strand cDNA was prepared with the primer adapter of sequence [5'
             GAGAGAGAGATTCTCGAGTTAATTAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved
             with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI.
             Host: SOLR.
FEATURES   Location/Qualifiers
     source   1..2135
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2900010C03"
                     /db_xref="MGI:1905833"
                     /db_xref="taxon:10090"
                     /clone="2900010C03"
                     /sex="male"
                     /tissue_type="hippocampus"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"

JOURNAL   prepare full-length cDNA libraries for rapid discovery of new genes
             Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED   11042159
REFERENCE   3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
             Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
             Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
             Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
             Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
             Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
             Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED   11076861
REFERENCE   4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
             FANTOM Consortium.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
             Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 4437)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
             Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
             Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
             Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
             Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
             Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
             Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
             Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
             Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
             Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
             Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..4437
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9930028B14"
                     /db_xref="MGI:2401399"
                     /db_xref="taxon:10090"
                     /clone="9930028B14"
                     /sex="female"
                     /tissue_type="vagina"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature   29..1968
                     /note="dystrophin, muscular dystrophy [MGD|MGI:94909,
                     GB|NM_007868, evidence: BLASTN, 100%, match=999)

misc_feature   1..2135
                /note="dystrophin, muscular dystrophy (MGD|MGI:94909,
                GB|NM_007868, evidence: BLASTN, 100%, match=999)"

RIGIN

Query Match   40.5%;   Score 526.4;   DB 11;   Length 2135;
Best Local Similarity   93.8%;   Pred. No. 2.2e-145;
Matches 548;   Conservative 0;   Mismatches 36;   Indels 0;   Gaps 0;

      718 CCACGAGACTCAAACAACAACTTGTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTT 777
          ||||| |||| |||| |||| ||| |||||||||||||||| |||||||||||||||||||
      144 CCACGAGACCCAAACCCACTTGTTGGGACCACCCCCAAAATGACAGAGCTCTACCAGTCTTT 203

      778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
          ||||||||||||| ||| |||||||| ||||||||||||||||||||||| ||||||||
      204 AGCTGACCTGAATAATGTCAGGTTCTCCGCGTATAGGACTGCCATGAAGCTCAGAAGGCT 263

      838 GCGAGAAGGCCCTTTGCTTGGGATCTCTTGAGCCTGCGCTGCCATGTGATGCCTTGGACCA 897
          |||| |||||||||| |||| ||| ||||||||||||| |||| ||| ||||| ||||||
      264 CCAGAAGGCCCTTTGCTTGGGATCTCTTGAGCCTGCGCTGCCATGTGATGCCTTGGACCA 323

      898 GCACAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATATTAATTGTTTGAC 957
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      324 GCACAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATAATTAACTGTTTGAC 383

      958 CACTATTTATGACCGGCCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTGCGT 1017
          ||| ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
      384 TACAATTTATGATCGTCGGAGCAAGAGCACAACAATCGGTCAACGTCCCTCTCTCGTGT 443

      1018 GGATATGTGTCTGAACTGCGCTGTGAAGTGTTATGATACGGGACGAACAGGGAGGATCCG 1077
          |||| ||| ||||||||| ||| ||||||| |||||||||||||||||||||||||||||
      444 GGATATGTGTCTCACTCGCTTCTCACTGGCTTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 503

      1078 TGTCCGTCTTTAAAACTGGCATCATTTCCCGTGTAAAGCACATTGGAAGACAAGTA 1137
          ||| ||| |||||||||||||| ||||| |||| ||||||||||||||||| ||||
      504 TGTCCGTCTTTTAAAACTGGCATCATTTCTCTGTGTGTAAAGCACACTGGAAGACAAGTA 563

      1138 CAGATACCTTTCAAGCAAGCAAGTGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGG 1197
          ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
      564 CAGATACCTTTCAAGCAAGCAAGTGCAAGTTCAACTGGCTTTGTGACCAGCGCAGGCTGGG 623

      1198 CCTCCTTCTGCATGCATTCATCCAAAATTCATCCAAATCCAAGACAGTTGGGTGAAGTTGGCATCCTTTGG 1257
          |||| ||||| ||| ||||||| |||||| |||| ||||||||||||||||||||||||||
      624 TCTTCTTCTGCATGCATTCATTCATTCAAAATCCATCCAAGACAGTTGGGTGAAGTTGCTTCCTTTGG 683

      1258 GGGCAGTAACATTGAGCCCAGTGTCCGGAGCTGCTTCCAATTG 1301
          ||||||||||||||||||||||||| |||||||||||||||||
      684 GGGCAGTAACATTGAGCCCAGTGTCCAGTTCAGGAGCTGCTTCCAATTG 727

RESULT 10
AK036936
LOCUS   AK036936   4437 bp   mRNA   linear   HTC 19-SEP-2003
DEFINITION Mus musculus adult female vagina cDNA, RIKEN full-length enriched
            library, clone:9930028B14 product:dystrophin, muscular dystrophy,
            full insert sequence.
ACCESSION   AK036936
VERSION   AK036936.1   GI:26085504
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE   High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED   10349636
REFERENCE   2
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to

```
polyA_signal    putative"
                4417..4422
                /note="putative"
polyA_site      4437
                /note="putative"

ORIGIN

Query Match      40.5%;   Score 526.4;  DB 11;  Length 4437;
Best Local Similarity  93.8%;  Pred. No. 3.8e-145;
Matches   548;  Conservative   0;  Mismatches  36;  Indels   0;  Gaps   0;

y     718 CCAGGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTT 777
          ||||||||| |||| ||||||||||||||||| |||||||||||||||||||||||||||||
b      49 CCACCGAGACCCCAAACCACTTGTTGGGACCACCCCCCAAAAATGACAGAGCTCTACCAGTCTTT 108

y     778 AGCTGACCTGAATAATGTCAGATTCCAGCTTGTATAGGACTGCCATGAAACTCCGAAGACT 837
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     109 AGCTGACCTGAATAATGTCAGGTTCAGGTTCTCCGGTATAGGACTGCCATGAAGAGCTCAGAAGGCT 168

y     838 GCAGAAGGCCCTTGCTTGGATCTCTTGAGCTGTCAGCTGCATGTGATGCCTTGGACCA 897
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     169 CCAGAAGGCCCTTGCTTGGATCTCTTGAGCTGTCAGCTGCATGTGATGCCCCTGGACCA 228

y     898 GCACAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATATTATTAATTGTTGAC 957
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     229 GCACAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATAATTAACTGTTTGAC 288

y     958 CACTATTTATGACCGCCTGGAGCAAGAGCACACAACAATTTGGTCAACGTCCCTCTGCGT 1017
          ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
b     289 TACAATTTATGATCGTCTGGAGCAAGAGCACACAACAATCTGGTCAATGTCCCTCTCTGTGT 348

y    1018 GGATATGTGTCTGAACTGGCTGCTGAAATGTTATGATACGGGACGAACAGGGAGGATCCG 1077
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     349 GGATATGTGTCTCCAACCGGCTTCTCAACCGGCTTCTCAATGTTATGATACGGGACGAACAGGGAGGATCCG 408

y    1078 TGTCCGTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACACATTTGGAAGACAAGTA 1137
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     409 TGTCCGTGCTTTTAAAACTGGCATCATTTCTCTGTGTAAAGCACACACTTGGAAGACAAGTA 468

y    1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGGCAGGCTGGG 1197
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     469 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGCCTTTTGTGACCAGGCGTAGGCTGGG 528

y    1198 CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGG 1257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     529 TCTTCTTCTGCATGATTCTATTCATTCAAATCCCCAAGACAGTTGGGTGAAGTTGCTTCTTTTGG 588

y    1258 GGGCAGTAACATTGAGCCGAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||
b     589 GGGCAGTAACATTGAGCCGAGTGTCAGGAGCTGCTTCCAATTTG 632
```

RESULT 11
AK075809
LOCUS   AK075809   1384 bp   mRNA   linear   HTC 20-SEP-2003
DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
            enriched library, clone:1810074E05 product:dystrophin, muscular
            dystrophy, full insert sequence.
ACCESSION   AK075809
VERSION   AK075809.1 GI:26344655
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus
ORGANISM   Mus musculus (house mouse)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE   High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED   10349636
REFERENCE   2
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

---

TITLE   Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED   11042159
REFERENCE   3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED   11076861
REFERENCE   4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1384)
AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Yasunishi,A.,
            Tejima,T., Toya,T., Yamamura,Y., Yamanaka,I., Yasunishi,A.,
            Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..1384
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:1810074E05"
                     /db_xref="MGI:1908112"
                     /db_xref="taxon:10090"
                     /clone="1810074E05"
                     /sex="male"
                     /tissue_type="pancreas"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="10 day old"
     CDS   120..1142

/note="unnamed protein product; dystrophin, muscular
dystrophy [MGD|MGI:94909, GB|NM_007868, evidence: BLASTN,
100%, match=999]
putative"
/codon_start=1
/protein_id="BAC35977.1"
/db_xref="GI:26344656"
/translation="MREHLKGHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKL
RRLQKALCLDLLSLSAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHNNLVN
VPLCVDMCLNWLLNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKRYLFKQVASSTGF
CDQRRLGLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIRAALFLDWMR
LEPQSMVWLPVLHRVAAAETAKHQAKCNICKECPIIGFRYRSLHGFHNYDICQGCFFSG
RVAKGHKMQHPYMVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLE
GDNMET"

ORIGIN

FEATURES   Location/Qualifiers
source   1..797
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:30344479"
                /tissue_type="Human Placenta"
                /lab_host="DH10B TonA"
                /clone_lib="NIH MGC 147"
                /note="Organ: placenta; Vector: pBluescriptR; Site_1:
                all-XhoI; Site_2: BamH; Oligo-dT primed using primer
                5'-TTTTTTTTTTTTTVN-3', size-selected for average
                insert size 2.3 kb and normalized to ROT 5. This is a
                primary library enriched for full-length clones and
                constructed using the Cap-trapper method (Carninci, in
                preparation). Library constructed by M. Brownstein
                (NIMH/NHGRI, National Institutes of Health). Note: This is
                a NIH_MGC library."

ORIGIN

Query Match   40.3%; Score 524.8; DB 11; Length 1384;
Best Local Similarity 93.7%; Pred. No. 4.8e-145;
Matches 547; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy   718 CCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT 777
          |||||||||||||||||||||||| |||||| |||||||||||||||||||||||||||||
Db   140 CCACGAGACCCAAACCACCACTTGTTGGGACCACCCCAAAATGACAGAGCTCTACCAGTCTTT 199

Qy   778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   200 AGCTGACCTGAATAATGTCAGGTTCTCCGCGATAGGACTGCCATGAAGCTCAGAAGGCT 259

Qy   838 GCAGAAGGCCCTTTGCTTGGATCTCTGAGCCTGTCAGCTGCTGCATGTGATGCCTTGGACCA 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   260 GCAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCCGCTGCATGTGATGCCCTGGACCA 319

Qy   898 GCACAACCTCAAGCAAATGACCAGCCCATGGATAATCCTGCAGATATCCTGAGATTATTAATTGTTTGAC 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   320 GCACAACCTCAAGCAAATGACCAGCCCATGGATAATCCTGCAGATAATCCTGCAGATAATTAACTGTTTGAC 379

Qy   958 CACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGT 1017
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   380 TACTAATTTATGACCGCGTCTGCTGGAGAGCAAGAGCACAACAATTTGGTCAATGTTTGATACGGG 439

Qy   1018 GGATATGTGTCTGAACTGCTGCTGCTGAATGTTTATGATACGGGACGGGACGAACAGGAGGGATCCG 1077
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   440 GGATATGTGTCTGAACTGGCTTCTCAATGTTTATGATACGGGACGGGACGGGACGAACAGGAGGGATCCG 499

Qy   1078 TGTCCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAGACAAGTA 1137
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   500 TGTCCTGTCTTTAAAACTGGCATCATTTCCTGTGTAAAGCACATTGGAGACAAGTA 559

Qy   1138 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCCAGGCTGGG 1197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   560 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCCAGGCTGGG 619

Qy   1198 -CCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGG 1239

RESULT 12
CB960722   797 bp   mRNA   linear   EST 29-APR-2003
DEFINITION AGENCOURT_13761995 NIH_MGC_147 Homo sapiens cDNA clone
            IMAGE:30344479 5', mRNA sequence.
ACCESSION   CB960722
VERSION   CB960722.1 GI:30216839
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

RESULT 13

```
LOCUS       AY408546    2874 bp    DNA    linear   GSS 15-DEC-2003
DEFINITION  Homo sapiens DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY408546
VERSION     AY408546.1  GI:39764517
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2874)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
  PUBMED    14671302
REFERENCE   2  (bases 1 to 2874)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..2874
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>2874
                     /gene="DRP2"
                     /locus_tag="HCM3260"
ORIGIN

Query Match    39.1%;  Score 508.2;  DB 29;  Length 2874;
Best Local Similarity  60.4%;
Matches  777;  Conservative  0;  Mismatches 510;  Indels  0;  Gaps  0;

Qy   14  GTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGAACCT  73
           || ||||| | ||||| || || ||| ||||  |  ||| |||| ||  || ||
Db  412  GTGCAACAGGAGGAGGAGACACATGCGGCCGTTGATGAAGAAGAGTCAAGTCTCGGGGCCCC 471

Qy   74  GTAATCATGATCATGTGCTGGAGTACTCTTGAGAGTGTACAGAATATTTCTGAAGATATTC 133
Db  472  TACATCTATTCTGTGCTGGAGTCAGCTCCAGAGCCCAGAGAGTGCGTCCTGCTCGGTG     531

Qy  134  CTAGAGAAACTCTACCAGGAGCCCGAGAGAGTCGGAAAGATACCTCCCCGAAACAGCCGGA 193
Db  532  TTAGGAGGAGCCCATTCTGATGAAGCAGAGGCGGCCACGCGTGGCCAGTGAACTG        591

Qy  194  ACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTG   253
Db  592  AGCCGCTTTGTATGAAGGAGGCCACGCGTGGCCAGTGAACTGGGGAGAAGTTGACAGCC   651

Qy  254  CACTCCCGCTGACTGGCAGAGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCAA 313
Db  652  CGCTGTGTGACCAGACCACCGTCACCATTGACGCGGACTCTTGGGAGCAGCTCTTCGGAGATTCAA   711

Qy  314  GAGGGCCACGGATGAGCTGACCTCAAGTCGCGCCAAGTGAGGTGATCAAGGGATCCTGG   373
Db  712  GGGGCAATGGAGGAACTAAGCACTACTACTCTGAGCAAGCTGAGGAGTTCCGAGCCACTTGG 771
```

```
                                           sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC
                                           3'].cDNA was cleaved with XhoI and BamHI. Vector: a
                                           modified pBluescript KS(+) after bulk excision from Lambda
                                           FLC I."

SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      ORIGIN
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 652)                                                     Query Match      38.6%;   Score 502.8;  DB 10;  Length 652;
AUTHORS      Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,            Best Local Similarity   92.6%;   Pred. No. 1.1e-138;
             Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,             Matches 539;   Conservative   0; Mismatches  42; Indels  1; Gaps   1;
             Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A.,Nomura,K.,
             Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
             Sano,H., Saaaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,              Qy  718 CCAGGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTT 777
             Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,                     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
             Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki.Y.        Db   45 CCACGAGACCCAAACCACTTGTTGGGACCACCCCAAAATGACAGAGCTCTACCAGTCTTT 104
TITLE        RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL      Unpublished (2001)                                                     Qy  778 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 837
COMMENT      Contact: Yoshihide Hayashizaki                                                |||||||||||||||||||||||||| ||||||||||||| |||||||||||||||||||||
             Laboratory for Genome Exploration Research Group, RIKEN Genomic       Db  105 AGCTGACCTGAATAATGTCAGGTTCTCCGGTATAGGACTGCCATGAAGCTCAGAGGGCT 164
             Sciences Center(GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)               Qy  838 GCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCCATGTGATGCCTTGGACCA 897
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
             Tel: 81-45-503-9222                                                   Db  165 CCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCCATGTGATGCCTTGGACCA 224
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.go.jp,                                    Qy  898 GCACAACCTCAAGCAAATGACCAGCCATGGATATCCTGCAGATTATTAATTGTTTGAC 957
             URL:http://genome.gsc.riken.go.jp/,                                           |||||||||||||||||||||||||||||||||||||||||| | |||||||||||||||
             Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,         Db  225 GCACAACCTCAAGCAAAATGACCAGCCATGGAATATCCTGCAGATAATTAACTGTTTGAC 284
             Itoh,M., Konno,H., Okazaki,Y., Muramatsu.M. and Hayashizaki.Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to        Qy  958 CACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCTCTGCGT 1017
             prepare full-length cDNA libraries for rapid discovery of new                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
             genes. Genome Res. . 10 (10), 1617-1630 (2000)                        Db  285 TACAATTTATGATCGTCTGGAGCAAGAGCACAACAATCTGTCAATGTCCCTCTCTGTGT 344
             wagi,K., Fujiwake,S., Inoue,K., Togawa,T., Izawa,M., Ohara,E.,
             Watahiki.M., Yoneda.Y., Ishikawa,T. Ozawa,K. Tanaka,T.,               Qy 1018 GGATATGTGTCGAACTGGCGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCG 1077
             Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.             ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
             and Hayashizaki.Y.                                                    Db  345 GGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 404
             RIKEN integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer. Genome Res.    Qy 1078 TGTCCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTA 1137
             10 (11), 1757-1771 (2000)                                                     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,             Db  405 TGTCCTGTCTTTTAAAACTGGCATCATTCCTGTGTGAAGCACACTTGGAAGACAAGTA 464
             Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA                 Qy 1138 CAGATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGCCAGGCTGGG 1197
             encyclopedia: real-time sequence clustering for construction of a            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
             nonredundant cDNA library. Genome Res . 11 (2), 281-289 (2001)        Db  465 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAGCGTAGGCTGGG 524
             Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A.,
             Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K.,         Qy 1198 CCCCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGCATCCTTTGG 1257
             Arakawa,T., Ishii,Y. and Hayashizaki.Y.                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
             Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.         Db  525 TCTTCTTCTGCATGATTCTATTCAAATCCCAAGACACGTGGTGAAGTTGCTTCCTTTGC 584
             Func. Genomics 2 pre, L72-L86 (2001
             Please visit our web site (http://genome.gsc.riken.go.jp) for         Qy 1258 GGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATT 1299
             further details.                                                              |||| |||||||||||||||||||||||||||| |||||||||
FEATURES     Location/Qualifiers                                                    Db  585 GGGCA-TAACATTGAGCCGAGTGTCAGGAGCTGTTTCAATTT 625
     source  1..652
             /organism="Mus musculus"
             /mol_type="mRNA"                                                      RESULT 15
             /db_xref="taxon:10090"                                                AY408548
             /clone="9930028B14"                                                   LOCUS       AY408548   2874 bp   DNA   linear   GSS 15-DEC-2003
             /sex="female"                                                         DEFINITION  Mus musculus DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
             /tissue_type="vagina"                                                             genomic survey sequence.
             /dev_stage="adult"                                                    ACCESSION   AY408548
             /lab_host="DH10B"                                                     VERSION     AY408548.1   GI:39764519
             /clone_lib="RIKEN full-length enriched, adult female                  KEYWORDS    GSS.
             vagina"                                                               SOURCE      Mus musculus (house mouse)
             /note="Site 1: SalI; Site 2: BamHI; cDNA library was                  ORGANISM    Mus musculus
             prepared and sequenced in Mouse Genome Encyclopedia                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Project of Genome Exploration Research Group in Riken                              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
             Genomic Sciences Center and Genome Science Laboratory in              REFERENCE   1  (bases 1 to 2874)
             RIKEN. Division of Experimental Animal Research in Riken              AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,.
             contributed to prepare mouse tissues. 1st strand cDNA was                         Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
             primed with a primer [5'                                                          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
             GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTTVN 3'], cDNA was                        Adams,M.D. and Cargill,M.
             prepared by using trehalose thermo-activated reverse                 TITLE       Inferring nonneutral evolution from human-chimp-mouse orthologous
             transcriptase and subsequently enriched for full-length by                       gene trios
             cap-trapper. cDNA went through one round of normalization             JOURNAL     Science 302 (5652), 1960-1963 (2003)
             to Rot = 10.0 and subtraction to Rot = 185.0. Second                  PUBMED      14671302
             strand cDNA was prepared with the primer adapter of                   REFERENCE   2  (bases 1 to 2874)
```

```
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
           Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2001) Celera Genomics, 45 West Gude Drive,
           Rockville, MD 20850, USA
COMMENT    This sequence was made by sequencing genomic exons and ordering
           them based on alignment.
FEATURES   Location/Qualifiers
source     1..2874
           /organism="Mus musculus"
           /mol_type="genomic DNA"
           /db_xref="taxon:10090"
gene       <1..>2874
           /gene="DRP2"
           /locus_tag="HCM3260"
ORIGIN

Query Match    38.2%;    Score 497;  DB 29;  Length 2874;
Best Local Similarity  59.8%;    Pred. No. 1.7e-136;
Matches   770;  Conservative   0;  Mismatches  517;  Indels   0;  Gaps   0;
```

```
Qy      14 GTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGAGGGAATTGAAAACTAAAGAACCT 73
           ||||||||| |||||     ||||||||||| |||||||||||||||||||||| |||||||
Db     412 GTACAACAGGGAAGGAGAGCACATGCAGCTTCAGGCTTCATGGAAGAAGTCAAGTCTAAGGGCCCC 471

Qy      74 GTAATCATGAGTACTCTTGAGACTTGTACGAAATATTTCTGACACAGAGCAGCCTTTGGAAGGA 133
           |||||||||||||||| ||||||||||||||| ||||||| |||||| ||||||| ||||||||
Db     472 TACATCTCCTCGTGCTTGAAATCCGCCAGGCTTTTCTGTCCCAGCATCCATTTGAAGAA 531

Qy     134 CTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTC 193
           |||||||| |||||||||||||||||||||| ||||| |||||||||||||||||||||||
Db     532 TTAGAGGAGTCTCAATCTGAGAGTAAAGATACCTCCCCAGACAGCAGCGGATTCAAAACCTT 591

Qy     194 ACTCCGCTTCTACGAAGCAGGCTGAGAGGTCAATACTGAGTGGGAAAAATTGAACCTG 253
           ||||| ||    ||||||||||| ||||||| |||   ||||||||||||||| |||||
Db     592 AGTCGCTTTGTATGGAAGCAGGCAACAGTGGCCAGTGGCCAGTGTGGGAGAAGCTGACAGCC 651

Qy     254 CACTCCGCTGACTGCGAGAGAAAATAGATGAGACCCTGAAAGACTTCCAGGAACTTCAA 313
           ||||| ||||| |||  |||||  ||||||||| ||||| |||||||||| |||||||||
Db     652 CGCTGTGTAGATCAGCATCGCCACATTGAGCATACTCTCTGGACAATCTATTGGAGATCCAA 711

Qy     314 GAGGGCCACGGATGAGCTGGACCTCAAGCTGCGCGCCAAGCTGAGCTGATCAAGGGATCCTGG 373
           ||||| |||||||||||||| |||||||| ||  |||||||||||||||| ||||||||||||
Db     712 GGGGGCAATGGAGGGAATTGAGCAGTACTTTGACCCCAAGCAGAGGGGAGTTCCGAGCCACGTGG 771

Qy     374 CAGCCCGTGGCGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCA 433
           |||||||||   ||||| |||||||||||||  |||||||| ||||| ||||||||||||||||
Db     772 GAGCCCATAGAGAGATCTTTATTGATTCCCTCCAAGAGAGCGTGAGCCACGCATATCAAGCCATCAAGCTA 831

Qy     434 CTTCGAGAGAAATTCGCGTCTGAAGAAGCAGCTGAGCCAGTCAATGACCTTGCTCGC 493
           ||||  ||||||||| | ||||||||||| || |||||||||||||| |||||||||
Db     832 TTCAAAGAAGAATTCTCTCCTCGTGAAAGATGGGGTGAAGTTAGTGAATGATCGGGCCCAC 891

Qy     494 CAGCTTACCACTTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTTCTGGGAAGACCTG 553
           ||||| |||| ||||||||||||  |||||||||  |  ||||||| ||||||||||||||
Db     892 CAGCTTGCCATCTCTGATGTGCACTTGTCAATGTGCACTTCTCAGGGCTCTCGGAAGACAGATC 951

Qy     554 AACACCAGATGAAGCTTCTGCAGAAGCTTCGCCAGGACGTGGCCGTCAGGCAGCTGCATGAA 613
           |||| ||| |||||||||||  |||| ||||||| |||||||||| |||||||||||||  ||
Db     952 AACATCCGGTGGAAACAGCTTCCAGGTGTCAGTGCTGCAGTGGAGGCTTAAGCAACTCCAGGAT 1011

Qy     614 GCCCACCAGGGACTTGGTCCAGCATCTGAAGCACTTTCTTTCCACGTCTGTCCAGGGTCCC 673
           |||||||||||||  |||||||||||| |||| ||| | |||||| |||||||||||||||
Db    1012 GCCCACCGGGGACTTTGGGCACTAGCAGTACTTTGGGCTTGCACAGTTCCCTCTGTCCAAGTTCCC 1071

Qy     674 TGGGAGAGCCATCTGCCAAACAAGTGCCCTACTATATCAACCACGAGACTCAAACA 733
           ||||||| |||   ||||  ||| | ||| ||||| ||||||||||||||||||||
Db    1072 TGGGAAAGAGCCAATTTCTCCCAATAAAGTCCCCTACTACATCAANNNNNNNNNNNN 1131

Qy     734 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 793
           || ||||| | |       |||||||||| |||||| ||| |||||||||||||  ||||
```

```
Db    1132 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCGATCTGAACAAC 1191

Qy     794 GTCAGATTCTCAGTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTGC 853
           | |||||||||||||| |||||||||||||| ||||||||||||||||||||||||||
Db    1192 ATTAAGTTCTCAGTTACCGCACTGCCATGAAGTCCGCAGAGTCCGCAGAAGGCCCTGCGC 1251

Qy     854 TTGGATCTCTTGAGCCTGTCAGCTGCATGTGCCTTGGACCAGCACAACCTCAAGCAA 913
           ||||||||||| |||||| |||| ||  |||||||||| ||    |||||||||||||
Db    1252 CTGGATCTGGTGAACTTAACTACAGCTCTGGGAGATCTTCAATGAGCATGACTTGCAGGCC 1311

Qy     914 AATGACCACCATGGATATCCTGCAGATTATTAATTGTTGACCACTATTATGACCGC 973
           |||||||||||| |||||| |||||| |||  ||||| | |||||| |||||||||
Db    1312 AGTGAACATGCATGATGGATGGGTGGAGTCATTCACTGCTTGACTGCCTTGTATGAACGA 1371

Qy     974 CTTGGAGCAAGAGACAACAATTTGGTCAACGTCCCCTTCTGCGTGATATGTGTCTGAAC 1033
           ||||| |||  || ||  ||||||  | ||||||||||| |   ||||| ||||||||||
Db    1372 CTGGGAGGAGGAAAGAGGCATCCCTGGTCAATGTGCCACTGTGTGTAGACATGAGCCTCAAC 1431

Qy    1034 TGGGCTGCTGAATGTTTATGATACGGACGAACAGAGGAGGATCCGTGTCCTGCTCTTTAAA 1093
           ||||| |||||||| || |||| |||||| |||||| |||||||||||||||||| |||||
Db    1432 TGGCTCCCCAATGTTTTTGATAGTGGTCGCAGTGGAAAGATGCGAGCATTGTTCCTTTTAAG 1491

Qy    1094 ACTGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAATTTGGAAGATACAGATACCTTTTCAAG 1153
           ||||||| ||| || |||||||||||||| |||||  || |||| ||| ||||||| ||||||||
Db    1492 ACTGGCATGCATGCCTGTCCGTGTGGCACGGAAGTGAAAGAAAAACTTCAGTATCTTCTTCAGC 1551

Qy    1154 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCGAGGCTGGGCCTTCCTCCTCTGCATGAT 1213
           |||||  | |||||||| |||  | |||| |||| |  || ||||| ||| ||| |||||||||
Db    1552 CAAGTAGCCAATTCAGGCGAGCAGCCAGTGTGATCACAGCCCATCTCCGGTGCCCGCTTCATGAA 1611

Qy    1214 TCTATCCAATTCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGCGGAGTAACATTGAG 1273
           |||| |||||| ||||||||| |||||||||| |||| |||||||| ||| || |||||
Db    1612 GCCATTCAAGTGCCCCGTCAGCTGGGTGAAGTGGCAGCAATTGGGGCAGCAATGTGGGAG 1671

Qy    1274 CCAAGTGTCCGGAGCTGCTTCCAATTT 1300
           |||||| |||||| ||||| ||||||
Db    1672 CCCAGTGTCCGTAGCGTGCTTTCGCTTT 1698
```

```
Search completed: April 5, 2004, 08:15:13
Job time : 2280.99 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2004, 16:58:13 ; Search time 97.0414 Seconds
(without alignments)
11443.127 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.........tcaaccacgagactcaaaca 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277745446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description	
1	1182.6	59.1	5952	4	US-09-687-875A-1	Sequence 1, Appli	
2	1182.6	59.1	13977	4	US-09-484-970B-60	Sequence 60, Appl	
3	1004	50.2	19307	3	US-08-836-022A-10	Sequence 10, Appl	c
4	1004	50.2	19307	4	US-09-427-048A-10	Sequence 10, Appl	c
5	397	19.8	6045	4	US-09-091-501B-7	Sequence 7, Appl	
6	397	19.8	10320	4	US-09-091-501B-9	Sequence 9, Appli	
7	311.8	15.6	3915	4	US-09-976-594-93	Sequence 93, Appl	
8	79.4	4.0	200	4	US-09-091-501B-5	Sequence 5, Appl	
9	78.6	3.9	200	4	US-09-091-501B-4	Sequence 4, Appli	
10	78.6	3.9	200	4	US-09-091-501B-6	Sequence 6, Appli	
11	76.6	3.8	7218	1	US-08-232-463-14	Sequence 14, Appl	c
12	63.6	3.2	238	4	US-09-687-875A-13	Sequence 13, Appl	
13	46.2	2.3	505	4	US-09-621-976-15639	Sequence 15639, A	c
14	44	2.2	1230025	4	US-09-198-452A-1	Sequence 1, Appli	c
15	43.4	2.2	1179	4	US-09-107-532A-1186	Sequence 1186, Ap	
16	42.8	2.1	1690	4	US-09-620-312D-69	Sequence 69, Appl	
17	42.8	2.1	7812	3	US-09-368-590-1	Sequence 1, Appli	
18	40.4	2.0	832	4	US-09-621-976-2813	Sequence 2813, Ap	
19	40.4	2.0	2223	1	US-08-257-073-4	Sequence 4, Appl	
20	39.2	2.0	16995	4	US-08-961-527-82	Sequence 82, Appl	
21	38.6	1.9	1751	4	US-09-620-312D-847	Sequence 847, App	
22	38.6	1.9	1995	1	US-09-425-069-3	Sequence 3, Appli	
23	38.6	1.9	1995	2	US-08-317-844B-3	Sequence 3, Appli	
24	38.4	1.9	7672	4	US-09-220-132-24	Sequence 24, Appl	
25	38.2	1.9	1131	6	5180810-3	Patent No. 5180810	
26	38.2	1.9	1784	6	5180810-2	Patent No. 5180810	
27	38	1.9	1394	4	US-09-247-155-76	Sequence 76, Appl	c
28	36.8	1.8	1886	6	5210183-1	Patent No. 5210183	
29	36.6	1.8	1845	4	US-08-887-534A-22	Sequence 22, Appl	
30	36.6	1.8	1845	4	US-09-527-431-22	Sequence 22, Appl	c
31	36.6	1.8	7075	3	US-08-956-171E-263	Sequence 263, App	
32	36.2	1.8	289	3	US-09-007-005-17	Sequence 17, Appl	
33	36.2	1.8	289	3	US-09-244-796-17	Sequence 17, Appl	
34	36.2	1.8	2447	2	US-09-014-969-14	Sequence 14, Appl	
35	36.2	1.8	11049	4	US-10-204-708-23	Sequence 23, Appl	c
36	36.2	1.8	168575	4	US-09-426-290-1	Sequence 1, Appli	
37	36	1.8	724	4	US-08-956-171E-832	Sequence 832, App	c
38	36	1.8	2873	4	US-08-630-915A-193	Sequence 193, App	
39	35.8	1.8	790	3	US-09-461-474-13	Sequence 13, Appl	
40	35.6	1.8	4868	1	US-08-139-937-12	Sequence 12, Appl	
41	35.6	1.8	4868	5	PCT-US93-11310-12	Sequence 12, Appl	
42	35.6	1.8	5934	4	US-09-418-710-2	Sequence 2, Appli	
43	35.6	1.8	8257	4	US-09-595-684B-30	Sequence 30, Appl	
44	35.6	1.8	8789	1	US-08-328-254-5	Sequence 5, Appli	
45	35.6	1.8	10136	1	US-08-353-700-2	Sequence 2, Appli	

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED P
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc_feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match   59.1%; Score 1182.6; DB 4; Length 5952;
Best Local Similarity   99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTGCCTG 872
        |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2946 GGAAGAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTGCCTG 3005

Qy 873 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 932
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3006 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 3065

Qy 933 CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3066 CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 3125

Qy 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAAAATCCTGAG 1052
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3126 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAAAATCCTGAG 3185

Qy 1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCTTGTTACAAAGACGTTTGGATAACATGAACTT 1112
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match          59.1%;  Score 1182.6;  DB 4;  Length 13977;
Best Local Similarity 99.7%;  Pred. No. 0;
Matches 1185;  Conservative 0;  Mismatches 4;  Indels 0;  Gaps 0;
```

Right-hand alignment blocks:

```
Qy    813 GGTACCTACTCATAGATTACTGCACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTG  872
          |||  | |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTG 8319

Qy    873 GCTTACAGAAGCTGAAACAACTGCCAATGTCTACAGGATGCTCCTACAGGATGCTACCCGTAAGGAAAGGCT  932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8320 GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTCCTACAGGATGCTACCCGTAAGGAAAGGCT 8379

Qy    933 CCTAGAGAACTCCAAGGAGTAAAGAGCTGATGAAACATGGCAAGACCTCCAAGGTGA  992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8380 CCTAGAGAACTCCAAGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8439

Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAG 1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8440 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAG 8499

Qy   1053 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTT 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8500 ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTT 8559

Qy   1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTAGGTCCCATTGGAGCCAGTTC 1172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8560 CAAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTAGGTCCCATTGGAGCCAGTTC 8619

Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTCTCCACCTTCTCCGCAGGACTTCGGTGGTGGCTGGTGGCTACAGCTGAA 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8620 TGACCAGTGGAAGCGTCTGCACCTTCTCCACCTTCTCCGCAGGACTTCGGTGGTGGCTGGTGGCTACAGCTGAA 8679

Qy   1233 AGATGATGAATTAAGCCGGCAGCGGCACCTATTGGAGCTCAGAGTTCAGAAGCA 1292
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8680 AGATGATGAATTAAGCCGGCAGCGGCACCTATTGGAGCTCAGAGTTCAGAAGCA 8739

Qy   1293 GAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8740 GAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 8799

Qy   1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 8859

Qy   1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAGTGGGAAAAATTGAACCTGCTGA 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8860 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAGTGGGAAAAATTGAACCTGCTGA 8919

Qy   1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8920 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 8979

Qy   1533 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGA 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8980 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGA 9039

Qy   1593 TGAGTGGGACCTCAAGCTGCGCGCCAAGCTGAGGTGATCAAGGATCCTGGCCAGCCCCGTGGG 1652
```

Left-hand alignment blocks:

```
Db   3186 ATCCCTGGAAGGTTCCGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTT 3245

Qy   1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTAGGTCCCATTGGAGCCAGTTC 1172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3246 CAAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTAGGTCCCATTGGAGCCAGTTC 3305

Qy   1173 TGACCAGTGGAAGCGCTGCGACCTTCTCGGTGTGGCTACAGCTGAA 1232
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3306 TGACCAGTGGAAGCGCTGCGACCTTCTCGGTGTGGCTACAGCTGAA 3365

Qy   1233 AGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAGCA 1292
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3366 AGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAGCA 3425

Qy   1293 GAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3426 GAACCATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 3485

Qy   1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3486 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT 3545

Qy   1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAGTGGGAAAAATTGAACCTGCTGA 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3546 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGAGTGGGAAAAATTGAACCTGCTGA 3605

Qy   1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3606 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 3665

Qy   1533 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGA 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3666 CTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGA 3725

Qy   1593 TGAGTGGGACCTCAAGCTGCGCGCCAAGCTGAGGTGATCAAGGATCCTGGCCAGCCCCGTGGG 1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3726 TGAGTGGGACCTCAAGCTGCGCGCCAAGCTGAGGTGATCAAGGATCCTGGCCAGCCCCGTGGG 3785

Qy   1653 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3786 CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGA 3845

Qy   1713 AATTGCGCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3846 AATTGCGCCCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3905

Qy   1773 TTTGGGCATTCAGCTCTACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 1832
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3906 TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG 3965

Qy   1833 GAAGCTTCTGCAGGTGGCCGTCAGGACGCTGCATGAGAGCCCACAGGGA 1892
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   3966 GAAGCTTCTGCAGGTGGCCGTCAGGACGCTGCATGAGAGCCCACAGGGA 4025

Qy   1893 CTTTGGTCCAGCATCCAGCCACTTCTTTTCCAGCTGTCCCAGGGTCCCTGGGAGAGAGC 1952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4026 CTTTGGTCCAGCATCCAGCCACTTCTTTTCCAGCTGTCCCAGGGTCCCTGGGAGAGAGC 4085

Qy   1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACCAGAGACTCAAACA 2001
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4086 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACCAGAGACTCAAACA 4134
```

```
RESULT 2
US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
```

```
Db  9040                                                          |||||||||||||||||||||
                      TGAGCTGGACCTCAAGCTGCGCGCCAAGCTGAGGTGATCAAGGAAAGTCAAGGACACTTCGAGCCCGTGGG  9099

Qy  1653  CGATCTCCTCATTGACTCTCGCGCCAAGATCACCTGAGGTGATCAAGGAAAGTCAAGGACACTTCGAGGAGA  1712
          |||||||||||||||||||||      |||||||||||      |||||||||||||||||||||||||||||
Db  9100  CGATCTCCTCATTGACTCTCTCCAAGATCACCTGAGAAAGTCAAGGACACTTCGAGGAGA  9159

Qy  1713  AATTGCGCCCTCGAAAGAGAGACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC  1772
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9160  AATTGCGCCCTCGAAAGAGAGACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC  9219

Qy  1773  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG  1832
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9220  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG  9279

Qy  1833  GAAGCTTCTGCGAGGTGGCCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA  1892
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9280  GAAGCTTCTGCGAGGTGGCCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA  9339

Qy  1893  CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTCGTCCAGGGTCCCTGGGAGAGAGC  1952
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9340  CTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCTCGTCCAGGGTCCCTGGGAGAGAGC  9399

Qy  1953  CATCTCGCCCAAACAAAGTGCCCTACTACTATATCAACCACGAGACTCAAACA  2001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9400  CATCTCGCCCAAACAAAGTGCCCTACTACTATATCAACCACGAGACTCAAACA  9448
```

US-08-836-022A-10

Query Match   50.2%;   Score 1004; DB 3;   Length 19307;
Best Local Similarity 90.3%;   Pred. No. 1.1e-302;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```
Qy   813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTG   872
          ||  |||||||||||||||||| ||||||||||| |||||||||| ||||||||  |||
Db  6434  GGAAGAAACTCATAGATTACTGCAGCAGTTCCCTGGACCTGGAGAAGTTCTTCTTCCTG  6375

Qy   873  GCTTACGAAGCTGAAACAACTGCCAATGTCCACAGATGCTACCCGTAAGGAAAAGGCT   932
          | |||||||||||||||||||||||||||| |||| ||     ||||||||||| ||||
Db  6374  GATTACGGAAGCAGAAACAACTGCCAATGTCCTACAGGACGCGTTCCCGTAAGGAGAAGCT  6315

Qy   933  CCTAGAAGACCTCCAAGGAGGTAAAGAGAGCTGATGAAGAACAATGGCAAGACCTCCAAGGTGA   992
          |||||||||||||||  |||||   ||||||||||||||  ||||||||||||| ||||||||
Db  6314  CCTAGAAGACCTCCAGGGAGTCAGAGAGCTGATGAAACCATGGCAAGATCTCCAAGGAGA  6255

Qy   993  AATTGAAGCTCACACAGATGTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAG  1052
          |||||||| ||||||||||| | |||||| | |||||||| | ||||||||||||||||||
Db  6254  AATTGAAACTCACACAGATATCTATCACACAATCTTGATGAAAATGCCAAAAAATCCTGAG  6195

Qy  1053  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAACACGTTTGGATAACATGAACTT  1112
          ||||||||||||||| ||||| ||    |||||||||| ||||||||||||||||| ||
Db  6194  ATCCCTGGAAGGTTCGGATGAAGCACCCCTGTTACAAGACGTTTGGATAACATGAATTT  6135

Qy  1113  CAAGTGGAGTGAACTTCAGATAACTTCAGATAAGTCTCTCAACATTAGTCCCATTGGAAGCAAGTTC  1172
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6134  CAAGTGGAGTGAACTTCAGATAACTTCAGATAAGTCTCTCAACATTAGTCCCATTGGAAGCAAGTTC  6075

Qy  1173  TGACCAGTGGAAGCCTCGCACCTTCTCGCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA  1232
          |||||||||||||  ||||  ||||| |  | |||||||||| ||| |||||||||||||
Db  6074  TGACCAGTGGAAGCGTTGCATCTTCTCTCTCTCTCAGAGAACTTCTTGTTGGTGGCTACAGCTGAA  6015

Qy  1233  AGATGATGAATTAAGCGCGAGGCACCTATTGGAGGGCACTTTCCAGCAGTTCAGAAGCA  1292
          ||||||||| |  ||| | || || | |  ||| ||  ||||| ||||||||||||||||
Db  6014  AGATGATGAACTGAGCCATGAGGCGACCCCATCGGTGTGATTTCCCAGCAGTTCAGAAGCA  5955

Qy  1293  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
          |||  || | ||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  5954  GAATGATATACATAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGAACCTGTAATCATGAG  5895

Qy  1353  TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT  1412
          |||||| |||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db  5894  TACTCTGGAGACTGTGAGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT  5835

Qy  1413  CTACCAGGAGCCCAGAGAGCTGCCTCCTGGAGGAGCCCAGATGTCACTCGGCTTCT  1472
          ||||||||||||||||||||||||||||||||| |||  || |||||||||||| ||
Db  5834  CTACCAGGAGCCCAGAGAGCTGCCTCCTGGAGAAGAGAGCTCAGGAATGTCACTCGGCTCCT  5775

Qy  1473  ACGAAAGCAGCAGCTGAGGAGTCAATACTGAGTGGGGAAAATTGAACCTGGGACTGA  1532
          |||||||||| |   |||| ||||  |  ||| || || ||||||||||||  |||||
Db  5774  ACGAAAGCAGGGTCAGGGTCAAGCTGAAGTGGGACAAATTGAACCTGGCTCAGTGA  5715

Qy  1533  CTGGCAGAGAAAAATAGATGATGAAGACCCTGAAAGACTTCAGGAACTTCAGAGAGGCCACGGA  1592
          ||||||||||||||||||||||||||| |||| |||||||||||||||||||||||| |||| ||
Db  5714  TTGGCAGAGAAAAATAGATGATGAAGACTCTTGAAAGACTTCAGGAACTTCAGAGAAGTGCCGGA  5655

Qy  1593  TGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCAAGGGCCGTGGG  1652
          ||| |||||||||| ||||||||||||||||||||||||| ||||||  ||||||
Db  5654  TGAACTGACCTCAAGTTGCGCCCAAGCTGAGGTGATCAAGGGATCAAGGCAGCCGTGGG  5595

Qy  1653  CGATCTCCTCATTGACTCTCTCCAAGATCACCTGTCTCGCCAGCTTACCAC  1712
          |||||||||||||||||||||||||||||||||    ||||||||||||||
Db  5594  GGATCTCCTCATTGACTCTCTCCAAGATCACCTGAAAAAGTCAAGGCACTTCGGGGAGA  5535

Qy  1713  AATTGCGCCCTCGAAAGAGAACCTGAGCGTCAATGACCTTGCTCGCCAGCTTACCAC  1772
          |||||  ||  |  ||||||||  |  | |||||||||| |||||||||||||||||||
Db  5534  AATTGACCTCTTAAAGAGAATGTCAATCGTGTCAATGACTTGCTCGCCAGCTTGACCAC  5475

Qy  1773  TTTGGGCATTCAGCTCTCACCCTTATAACCTCAGCACTTTGGAAGATCTGAATACCAGATG  1832
          | ||| ||||||||||||| |||||||||||||||||||| ||||| |||||||||||||||
Db  5474  ACTGGGCCATTCAGCTCTCCACCCTTATAACCTCAGCACTTTGGAAGATCTGAATACCAGATG  5415
```

RESULT 3
US-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
;; APPLICANT: Trustees of the University of Pennsylvania
;; APPLICANT: Wilson, James M.
;; APPLICANT: Fisher, Krishna J.
;; APPLICANT: Chen, Shu-Jen
;; APPLICANT: Weitzman, Matthew
;; TITLE OF INVENTION: Improved Adenovirus Virus and
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;;   ADDRESSEE: Howson and Howson
;;   STREET: Spring House Corporate Cntr, P O Box 457
;;   CITY: Spring House
;;   STATE: Pennsylvania
;;   COUNTRY: USA
;;   ZIP: 19477
;; COMPUTER READABLE FORM:
;;   MEDIUM TYPE: Floppy disk
;;   COMPUTER: IBM PC compatible
;;   OPERATING SYSTEM: PC-DOS/MS-DOS
;;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;;   APPLICATION NUMBER: US/08/836,022A
;;   FILING DATE:
;;   CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;;   APPLICATION NUMBER: US 08/331,381
;;   FILING DATE: 28-OCT-1994
;; ATTORNEY/AGENT INFORMATION:
;;   NAME: Bak, Mary E.
;;   REGISTRATION NUMBER: 31,215
;;   REFERENCE/DOCKET NUMBER: GNVPN.008PCT
;; TELECOMMUNICATION INFORMATION:
;;   TELEPHONE: 215-540-9200
;;   TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;;   LENGTH: 19307 base pairs
;;   TYPE: nucleic acid
;;   STRANDEDNESS: double
;;   TOPOLOGY: unknown
;;   MOLECULE TYPE: cDNA

```
Db   6374 GATTACGGAGCAGAAACAACTGCCAATGTCCTCAGGACGCGTTCCCGTAAGGAGAAGCT 6315
                                                                      
Qy    933 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGAACCTCAAGGTGA 992
          ||||||   ||||||  ||||    ||||| ||||||||  ||||  ||  |||||| ||
Db   6314 CCTAGAGAGACTCCAGGGAGTCAGAGAGCTCAGAGAGCTGATGAAACCATGCAAGAGATCTCCAAGGAGA 6255

Qy    993 AATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAG 1052
          |||||| | |||||||||| | |||||| | ||||||||  |||||||||||||||||||
Db   6254 AATTGAAACTCACACAGATCTATCACAATCTTGATGAAAATGGCCAAAAAATCCTGAG 6195

Qy   1053 ATCCCTGGAAGGTTCGGATGAGCAGTCCTGTTACAAGACGTTTGGATAACAATGAACTT 1112
          ||||||||||||||||||||||||  |   ||||||| ||||||||||| |||||| ||
Db   6194 ATCCCTGGAAGGTTCGGATGAAGCACACCCCGTTACAAAGACGTTTGGATAACAATGAATTT 6135

Qy   1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACAATATAGTCTCCAACATTGGAAGCCAGTTC 1172
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db   6134 CAAGTGGAGTGAACTTCAGAAAAAGTCTCTCAACAATATAGTCTCCAACATTGGAAGCAAGTTC 6075

Qy   1173 TGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGTGTGGCTACAGCTGAA 1232
          |||||||||||||||| ||||| |||||||||||||||||| ||| ||||||||||||||
Db   6074 TGACCAGTGGAAGCGTTGCATCTTTCTCTGCAGGAACTTCTTGTTGTTGGCTACAGCTGAA 6015

Qy   1233 AGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA 1292
          |||||||||| | |||||| ||||||||   | |||| | |||||||||||||||||||||
Db   6014 AGATGATGAACTGAGCCGTCAGGCACCATCGGTGTGATTTCCCAGCAGTTCAGAAGCA 5955

Qy   1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTCAAAGAACCTGTAATCATGAG 1352
          |||| ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5954 GAATGATATACATAGGGCCTTCAAGAGGGAATTGAAAACTCAAAGAACCTGTAATCATGAG 5895

Qy   1353 TACTCTTGAGACTGTACGAATTATTCTGACAGAGTGGGGGAAAGACCTAGAGAAACT 1412
          ||||||||||||||| | ||| ||||||||||| |   ||||||||| || ||||||
Db   5894 TACTCTTGAGACTGTGAATATTTCTGACAGAGACCAGCAGCCTTTGGAGGACCTAGAGAAACT 5835

Qy   1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGTGCCAGAATGTCCAGTCCACTCGGCTTCT 1472
          |||||||||||||||||||||||||||||||||||   ||||||||||| || || ||| |||||
Db   5834 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAAAGAGCTCAGAGATGTCAGCTCACTCGGCTCCT 5775

Qy   1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGGAAAAATTGAACCTGACACTCGCTGA 1532
          ||||||||||||||||||||||| | ||||||| |||| |||||||||||||||| |||||
Db   5774 ACGAAAGCAGGCTGAGGAGGTCAACGGTCAATGGGGAATGGGAACAATTGAACCTGACACTCGCTGA 5715

Qy   1533 CTGGCAGAGAAAATAGATGAGACCCTGAAAGATCTCCAGGAACTTCAGGAGGTGCACGGA 1592
          |||| |||||||| ||||||| |||||||||||| |||||||||||||||||||| ||||
Db   5714 TTGGCAGAGAAAAATAGATGAGCTCTTGAAAGATCCCCAGGAACTTCAGGAAGTGCCCGA 5655

Qy   1593 TGAGCTGCACCTCAAGCTGCGCGCAAGCTGAGGTGATCAAGGATCCTCGGCCCCGTGGG 1652
          ||| |||||||| ||||| ||||| ||||| ||||||||||||||| ||| ||||||| 
Db   5654 TGAACTGGACCTCAAGTTGCGCGCAAGCTGCGCGTGATCAAGGATCCTCGGCAGCCAGTGGG 5595

Qy   1653 CGATCTCCTCATTGACTCTCTCCAAGATCCTCCAGATCACCTCGAGCAAGTCACTTCAGGGAGA 1712
          ||||| |||||||||| ||||| ||| ||  |||||||||||| || |||||||||| ||||||
Db   5594 GGATCTCCTCATTGACTCTCTCCAAGATCACCTTGAAAAAGTCAAGCAAGCACTTCGGGGAGA 5535

Qy   1713 AATTGCCGCTCCTCGGAAAGAGAACGAACGTGAGCACAGTCATGACCTTGTCTCGCCAGCTTACCAC 1772
          ||||| | |||||  ||||||||| | ||||| |||  ||||||||||||||||||| ||||||||||
Db   5534 AATTGCACCTCCTCTAAAGAGAATGTCAATCGTGTCAATGACCTTGACACTTCAGTCAGCTGACCAC 5475

Qy   1773 TTTGGCATTCAGTCTCCACCGTATAACCTCACCCTTATACCTACAGGACCACGAACACCAGATG 1832
          |||||||||||||| |||||| || |||| ||||| || |||| || ||||||||||||||||
Db   5474 ACTGGGCATTCAGTCTCAGTCTCTTTAAAGATCTGAATACCACCAGATG 5415

Qy   1833 GAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGA 1892
          |||||||||| ||||||| | ||  ||||| ||||| |||||||||||||||||||||||
Db   5414 GAGGCTTCTACAGGTGGCTGCGTGTGGAGGACCGTGTCAGACAGCTGCATGAAGCCCACAGGGA 5355

Qy   1893 CTTTGGTCCAGCATCCAGCACTTTTTCCACGTCTGTCCAGGGTCCCTGGACCTGGAGAGC 1952
          |||||||| || ||||||||||||| ||| ||  || ||  |||||||||| ||||||||
Db   5354 CTTTGGTCTGCGATCCCCAGCCACTTCCTTTCCACTTCAGTTCCCTCGGGTCCCTGGGGAGAGC 5295

Qy   1953 CATCTCGGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 2000
          |||||| |||||||||||||||||||||| |||||||||||||||||||
Db   5294 CATCTCACCCAAACAAAGTGCCCTACTATATCAACCACGAGACCCAAAC 5247
```

RESULT 4
US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
;   Wilson, James M.
;   Fisher, Krishna J.
;   Chen, Shu-Jen
;   Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
;   Methods of Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Howson and Howson
;   STREET: Spring House Corporate Cntr, P O Box 457
;   CITY: Spring House
;   STATE: Pennsylvania
;   COUNTRY: USA
;   ZIP: 19477
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/427,048A
;   FILING DATE: 21-Oct-1999
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/836,022
;   FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Bak, Mary E.
;   REGISTRATION NUMBER: 31,215
;   REFERENCE/DOCKET NUMBER: GNVPN.008PCT
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 215-540-9200
;   TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 19307 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: unknown
;   MOLECULE TYPE: cDNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-09-427-048A-10

```
Query Match           50.2%;   Score 1004;  DB 3;  Length 19307;
Best Local Similarity 90.3%;   Pred. No. 1.1e-302;
Matches 1073;  Conservative    0;  Mismatches 115;  Indels    0;  Gaps    0;

Qy    813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAGTTTCTTGCCTG 872
          ||  |  |||||||||||||||||| ||||||| ||||||||| |||||||||| |||
Db   6434 GGAAGAAACTCATAGATTACTGCAGCAGTTCCCTCTGGACCTGGGAAGTTTCTTCCTG 6375

Qy    873 GCTTACAGAAGCTGAAACAACTGCCAATGCTCCTACAGGATGCTACCCGTAGGAAAGGCT 932
          | || |||||  |||||||||||||||||| | ||||||||||| ||||||||||||| |
```

RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match   19.8%; Score 397; DB 4; Length 6045;
Best Local Similarity 59.3%; Pred. No. 4.2e-113;
Matches 698; Conservative   0; Mismatches 470; Indels   9; Gaps   1;

Qy 833 TGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAA 892
        |||| ||||| |||||||||| |||| || || |||||||| ||||||||||||||
Db 3069 TGCAGGCCTCTCGCGAGAGATCTGGAAAACTTCCTGAAGTGGATCCAAGAAGCAGAGACCA 3128

Qy 893 CTGCCAATGTCCTACAGAGATGCTACCCGTAAGGAAAAGCTCCTAGAAGACTCCAAGGGAG 952
        |||| | ||||||||| | || ||||| |||||||||| ||| ||||| || ||||||
Db 3129 CAGTGAATGTGCTTGTGATGCCTCTCATCGGGAGAATGCTCTTCAGGATAGTATCTTGG 3188

Qy 953 TAAAGAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAAGTTGAAGCTCACACAGATG 1012
        |||| |||| |||||| | ||| || |||||||||||| ||| |||||||| || ||
Db 3189 CCAGGGAACTCAAACAGCCAGATGCAGCATCCAGGGACAGAAATTGATGCCCACAAATGACA 3248

Qy 1013 TTTATCACAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATG 1072
        |||| |||||| ||||| ||| || |||| ||| |||||| |||| |||||| |||||||
Db 3249 TATTTAAAAGCATTGACGGAAACAGGCAGAGATGGTAAAGCTTTGGGAAATTCTGAAG 3308

Qy 1073 ATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGAGTGAACTTCGGA 1132
        |||| ||| |||||||||| || ||||| ||||| ||||||| |||| |||||| ||
Db 3309 AGGCTACTATGCTTCAAACATCGACTCGACTGGATGGATGATATGAACCAAAGATGACTTAAAAAG 3368

Qy 1133 AAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGC 1192
        |||| | ||||| |||||| |||| ||||||| ||| |||| ||||| |||||||||||
Db 3369 CAAAATCTGCTAGCATCAGGCCCCATTTGGAGGCCTAGCTGGCTTGCGTGAGAAGATGGAACAGGTTGC 3428

Qy 1193 ACCTTTCTCTGCGAGGAACTTCTCGGTGTGGCTACAGCTGAAAATGATGAATTAAGCCGGC 1252
        |||| ||| || |||| |||||| ||||||| |||||| ||| |||||| ||||| ||
Db 3429 TGAGTGCCTTAGAGAACTGATCAAATGCTAAATATGAAAGATGAAGAGCTTAAGAAAC 3488

Qy 1253 AGGACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGATGTACATAGGGCCT 1312
        ||||| |||||||||| || ||| |||||||| ||| || |||| |||||| |||
Db 3489 AAATGCCCATTGGAGGAGATGTTCCAGCTCCAGCCTTACAGCTCCAGTATGACCATTGTAAGGCCC 3548

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320

Qy 1313 TCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAA 1372
        ||||| ||| ||||||||||||||| |||||| |||||| ||||| |||||||||| |||
Db 3549 TGAGACGGGAGTTAAAGGAGCAGCCTTT------GGAAGGACTAGAGAAACTCTACCAGGAGC 3608

Qy 1373 TATTTCTGACAGAGCAGCCTTT------GGAAGGACTAGAGAAACTCTACCAGGAGC 1423
        ||||||||||||||||||||| |||||||||||||||||||||||||||||
Db 3609 TTTTCTTGGCTGATCAGCCAATTGAGGGCCCCTGAGAGCCAAGAGAGAAACCTACAATCAA 3668

Qy 1424 CCAGAGAGCTGCCTCCTGAGGAGAGAGGCCCAGAATGTCACTCGGCTTCACGAAAGCAGG 1483
        |||||||| ||| || |||||||||||||| |||| ||||||||| ||||| || ||
Db 3669 AAACAGAATTAACTCCTGAGGAGAGAGGCGCCAAAAGATTGCCAAAGCCATGCGCAAACAGT 3728

Qy 1484 CTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCCACTCCGTGACTGGCAGAGAA 1543
        ||||| || ||||| ||||||||||||| ||||| |||| | ||| |||||||| |||||
Db 3729 CTTCTGAAGTCAAAGAAAAAATGGAAAGTCTAAATGCTGTAACTAGCAATTGGCAAAAGC 3788

Qy 1544 AAATAGATGAGACCCTTGAAAGACTTCAGAGACTTCAAGAGCCACGGATGAGCTGGACC 1603
        |||| ||||||| || |||||||||||||| |||| |||||| |||| ||||||||||
Db 3789 AAGTGGACAAGGCATTGGAGAAACTCAGAGACCTGCACAGGAGCATGGATGGACCTGGACG 3848

Qy 1604 TCAAGCTGCGCCAAGCTGGAGTGAGTGATCAAGGAGATCCTGGCACGCCCCGTGGGCGGAT 1663
        |||||||| | ||||||||| ||||| |||||||||| ||||| |||| || |||| |||
Db 3849 CTGACATGAAGGAGGCAGAGTCCGTGCGGAATGGCTGGAAGCCGCCGGTGGGAGGACTTACTCA 3908

Qy 1664 TTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCCACTTCGAGAGAGAAATTGGGCCTC 1723
        |||||||| |||||| ||||||||||||| |||| ||||||||||||| ||||| || ||
Db 3909 TTGACTCGCTGCAGGATCACATTGAAAAATCATGGCATTTAGAGAAGAAATTGCACCAA 3968

Qy 1724 TGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTC 1783
        |||| ||| ||| ||| | ||| |||||| ||||| |||||||| |||| ||||| ||||
Db 3969 TCAACTTTAAAGTTAAAACGGTGAATGATTTATCCAGTCAGTCGTCTCCACTTGACCTGC 4028

Qy 1784 AGTCTCCACCGTATAACCTCAGCACTCTGGAAGACTTGAACACCAGATGGAAGCTTCTGC 1843
        || |||| | ||||| ||||| |||| |||||| || ||||||| |||||||| ||||||
Db 4029 ATCCCTCTCTAAAGATGTCTCGCCAGCTAGATGACCTTAGATGACCTGATGGAAACTTTTAC 4088

Qy 1844 AGGTGGCCGTGCAGGACCGAGTCAGGCAGCGTGCATGAGAGCCACAGAGCTTGGTCCAG 1903
        ||||| |||||||||||| |||||||| ||| |||||| ||||||||||||| |||||
Db 4089 AGGTTTCTGTGATGATCGCCTTAAACAGCTTCAGGAAGCCCACAGAGATTTGACCAT 4148

Qy 1904 CATCTCAGCACTTTCTTCCACGTCTGTCCACGGGTCCCTGGGAGGAGCCATCTCGCCAA 1963
        |||||||||||||| ||||||||| || ||||||| |||||| ||||| |||||||||
Db 4149 CCTCTCAGCATTTTTCTCTCTACGTCAGTCCAGTCCAGTCCGGCGTGCCAAGATCCATTTCACATA 4208

Qy 1964 ACAAAGTGCCCTACTACTATATCAACCACCAGAGACTCAAAC 2000
        ||||||||| ||||||| || || |||||||| ||||||
Db 4209 ATAAAGTGCCCTATTACATCAACCATCAAACACAGAC 4245

```
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (11)..(10312)
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Full length
;   OTHER INFORMATION: utrophin construct
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (724)..(758)
;   OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match           19.8%;  Score 397;  DB 4;  Length 10320;
Best Local Similarity 59.3%;  Pred. No. 6.2e-113;
Matches  698;  Conservative    0;  Mismatches  470;  Indels    9;  Gaps    1;

Qy    833 TGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGCTGAAACAA  892
              ||| ||||||| |||| | |||||| || |  ||||| | ||| || ||||| |||
Db   7344 TGCAGGCCTCTCGCAGAGATCTGGAAAACTTCCTGAAGTGGATCCAAGAGACAGAGACCA 7403

Qy    893 CTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAAGGCTCCTAGAAGACTCCAAGGGAG  952
              ||| || | || ||| ||||  ||||| | || ||| |  |||| ||||| |||||
Db   7404 CAGTGAATGTGCTTGTGGGATGCCTCTCATCGGGAGAATGCTCTCAGGATAGTATCTTGG 7463

Qy    953 TAAAAGAGTCGATGAAACAATGGCAGAGACCTCCAAGGTGAAAATTGAAGCTCACACAGATG 1012
              ||||  ||| || ||||  ||||||  || ||  |||| ||| |||||| || |||| ||
Db   7464 CCAGGGAACTCAAACAGCAGATGAGAACAGCCAAAAATCCAGGCAGAATTGATGCCCACAATGACA 7523

Qy   1013 TTTATCACACCTGGATGAAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCCGATG 1072
              ||||  | |  | |||||  || || |||| |||| | ||| |||||  | || |||
Db   7524 TATTTAAAAGCATTGACGAAGACAGGCAGAGGATGGTAAAGCTTGGGAAAAATTCTGAAG 7583

Qy   1073 ATGCAGTCCGTTACAAAGACGTTTGGATAACATGATACTTCAAGTGAGTGAACTTCCGA 1132
              ||||| || ||| ||| ||  |||||| |  | ||  ||||| || ||| ||||| ||
Db   7584 AGGCTACTATGCTTCAACATCGACTGGATGATATGAACCAAAGACAGACTTAAAAG 7643

Qy   1133 AAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCGACCAGTGGAAGCGTCTCG 1192
              ||||| |||  || || ||  |||  || |||| ||||||| ||| || |||| || |
Db   7644 CAAAATCTGCTAGCATCAGGGCGCTAGCCCATTTGGAGGCCGCAGCGTGGAACCAGGTTGC 7703

Qy   1193 ACCTTTCTCTGCAGGAACTTCTGGTTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGC 1252
              |||||  | ||||  | |||||||   | || | | |  ||| ||||| ||  | |||||
Db   7704 TGATGTCCTTAGAAGAACTGAATCAAATGGCTCAAATGCTGAATATGGATAGACAGAGAAAC 7763

Qy   1253 AGGCCACCTATTGGAGCGACTTCCAGAGTTCAGAAGCCAGAGACGACGATGTACATAGGGGCT 1312
              |||||  || ||||||  | ||||  | | ||||||| |||  | | || | | ||||| |
Db   7764 AAATGCCATTGGGAGGAGATGTTCCAGCTCACAGTCCTTACAGCTCCAGTATGACCATTGTAAGGCCC 7823

Qy   1313 TCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTAAGAA 1372
              ||| ||| |  |||  | | ||| |  ||  || |||||| | | ||||||  || || ||
Db   7824 TGAGACCGGGAGTTAAAGGAGTAAGGAGAATATTCTGTCCTGAATGCTGTCGACCAGGCCCAG 7883

Qy   1373 TATTTCTGACAGAGCAGCCTTT------GGAGCGACTAGAGAAACTCTACCAGGAGC 1423
              | | |  | || |||||  |||       |||  | |||||| | || |  | ||
Db   7884 TTTTCTTGGCTGATCAGCCAATTGAGGGCCCCTGAGAGCGCAAGAGAACCTACAATCAA 7943

Qy   1424 CCAGAGAGCTGCCTCCTGAGGAGAGAGCCCCAGAAATTGAACCTTCACGAAAGCAGG 1483
              || ||| | | ||||| ||||| || || |||| | |||||   ||| |||| |
Db   7944 AAACAGAATTAACTCCTGAGGAGAGCAGCAGCCCAAAAGATTGCCGCAAAGAGCCAAACAGT 8003

Qy   1484 CTGAGGAGGTCAATACTGAGTGGGAGAAAATTGAACCTGCGACTCCGCTGACTGGCAGAGAA 1543
              || |||||| | || | ||| |||| ||||| |||| ||  || | | ||| ||| || |
Db   8004 CAGCCAAAAGATGAATTAAGGCGGGACAGAAGTGTCAGCTCAGTTTATACTAGCAATTGGCAAAAGC 8063

Qy   1544 AAATAGATGAGACCCTGAAAGACTTCCAGAGAACTTCCAAGAGGCCACGGATGAGCTGGACC 1603
              ||||  || |   | |||||| | || ||| || |  |   ||||| ||| || |||| ||
Db   8064 AAGTGGACAAGGCATTGGAGAACTTCAGGAGACCTGCAGGGAGCTATGGATGACGACTGGACG 8123

Qy   1604 TCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTCTGGTCCGCGGCGATTCCTCA 1663
              |||| || | || ||| ||||| | |  ||||| || ||  | |||| || |||| ||
Db   8124 CTGACATGAGTACTCTTGAGACTGTACGAATGATATTCTGACGAGACTATTTCTGCCAGGCCTTCCTCCA 8183
```

```
Qy   1664 TTGACTCTCTCCAAGATCACCTCCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTC 1723
              |||||||||| ||||| |||||| |||||||  | |||| || ||||||||| ||||||
Db   8184 TTGACTCCGCTGCAGGATCACATTGGATCGCTTCCGCCAGCTTACCACTTTGGGGCATTC 8243

Qy   1724 TGAAAGAGACGTGAGCCACGTCAATGACTTGCTCGCCAGCTTACAGCTTTTGGGCATTC 1783
              | |||| || ||  |||| || ||  || || | || || | ||| | || |||||||
Db   8244 TCAACTTAAAGTTAAAACGGTGAATGATTATTCCAGTCAGTGTCTCCACTTGACCTGC 8303

Qy   1784 AGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGGAAACACCAGATGGAAGCTTCTGC 1843
              |||| |||     | ||  || || |  | | ||   | | || || |  || | | |||||
Db   8304 ATCCCTCCTCTAAAGATGTCTCGCCAGCTAGATGACCTTAATATGCGATGGAAACTTTTAC 8363

Qy   1844 AGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAGCGCCACAGGGACTTTGGTCCAG 1903
              |||| |  | || ||  || | |   ||  ||||  | ||| |||| |||| |||||  ||
Db   8364 AGGTTTCTGTGGATGATCGCTTAAACAGCTTCAGGAAGCCCCACAGAGATTTTGGACCAT 8423

Qy   1904 CATCTCCAGCACTTTCTTCTTCCCACGTCGTGTCCAGGGTCCTGGAGAGGGCCATCTCGCCAA 1963
              |  ||| |  |||||| |  |  | || |||| ||  || ||  |||  | |  || ||||||
Db   8424 CCTCTCAGCATTTTCTCTTCTTCTTACGGTCAGTCCAGTGGTGCCGCGAAAGATCCATTTCACATA 8483

Qy   1964 ACAAAGTGCCCTACTATCAACCACAGACTCAAAC 2000
              ||  |||||| ||  | ||| | |||||||  |
Db   8484 ATAAAGTGCCCTATTACATCAACCATCAACACACAGAC 8520
```

```
RESULT 7
US-09-976-594-93
;   Sequence 93, Application US/09976594
;   Patent No. 6673549
;   GENERAL INFORMATION:
;   APPLICANT: Furness, Michael
;   APPLICANT: Buchbinder, Jenny
;   TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;   FILE REFERENCE: PA-0041 US
;   CURRENT APPLICATION NUMBER: US/09/976,594
;   CURRENT FILING DATE: 2001-10-12
;   PRIOR APPLICATION NUMBER: 60/240,409
;   PRIOR FILING DATE: 2000-10-12
;   NUMBER OF SEQ ID NOS: 1143
;   SOFTWARE: PERL Program
;   SEQ ID NO 93
;   LENGTH: 3915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 290344.1
US-09-976-594-93

Query Match           15.6%;  Score 311.8;  DB 4;  Length 3915;
Best Local Similarity 60.0%;  Pred. No. 1.4e-86;
Matches  538;  Conservative    0;  Mismatches  357;  Indels    2;  Gaps    1;

Qy   1104 CATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCCAACATTAGGTCCCATTTGGA 1163
              |||||||||||||| |||||||||||||||||||||||| | ||||||||||||  |
Db    544 CATGAATCTGTGTTGGAATGAAATAAAAAAAAGTCTCCACAACCTCCGCCTCGCTAGA 603

Qy   1164 AGCCAGTTCGACCAGTGGAAGCGTCTGACCAGTTGGAAGACTTCTGTGTGGCT 1223
              ||||| || | | |  || |||||| |||  |  || |   || | |||||||
Db    604 GGCCTTCTCAGCCACCACCAGCTTCAGCTCCAGCTTCAAGAGATTATTGACTGGCT 663

Qy   1224 ACAGCTGAAAGATGATGAATTAAGCCGGCACTTTGGAGGCGACTTTCCAGCAGT 1283
              |||| ||  | |||||| | |||  |  || |  |||| |||| |  ||  | |||
Db    664 CAGCCAAAAGCATGAAGATGAGAGGAGTTGTCAGCTCAGTTTATGCTAGCAATTGGCCCTTGT 723

Qy   1284 TCAGAAGCAGAACGATGTACATAGGGGATTGAAAACTAAAGAACCTGT 1343
              || ||| ||| || || ||| || |  |||  | | | ||| | ||||
Db    724 GCAACAGGAGAAGGAGACACATGCGGCCATGAGGATGAGTCAAGTCTCCGGCCCCCTA 783

Qy   1344 AATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTCCTCA 1403
              | |||||||| | |||||||||  |||||||  |  | || |||| || || | |
Db    784 CATCTATTCTGTCCTGAGTGTGAGCTCAGTCAGCTGGAGTAGTCCAGCACCCCATTTGAGGAGTT 843
```

```
Qy    1404  AGAGAAACTTACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCAC  1463
                  ||||||  |||| |||||||||| |||| ||  | | |||||| ||  ||||||||||
Db     844  AGAGGAGCCCTCATTCTGAGAGCAGGGTCAAAGATACCTCCCGAAAACAGCGGATCCAGAATCTCAG  903

Qy    1464  TCGGCTTCACGAAAGCAGGAGCAGCGGCACGGTGCCCAGTGAACTGTGGGAGAAGTTGACACCTGCA  1523
                  |||||| |||  ||| ||||||||| || | ||||||| |||||||  ||| |||||||
Db     904  CCGCTTTGTATGGAAGCAGCAGGCGACGGACCGTGCCCAGTGAACTGTGGGAGAAGTTGACAGCCCG  963

Qy    1524  CTCCGCTGACTGACGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGA  1583
                  ||||||||  ||  ||||| ||||||  || || ||||| |||||| |||||||||||| ||
Db     964  CTGTGTGGACCAGCACCGTCACATTGAGCGGACTCTGGAGCAGCTCTTGTGGGAGATTCA--G  1021

Qy    1584  GGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCA  1643
                  ||||| ||||||||| |||||||||||||| ||| | | ||||||| |||||||||||
Db    1022  GGGCATGGAGGAACTAAGCACTACTCTGAGCCGAGGTCGAGCCACTTGGGA  1081

Qy    1644  GCCCGTGGGCGCGATCCTCCTCATTGACTCTCTTCCAGATCACTCGAGAAAGTCAAGGCACT  1703
                  ||||| || || || |||||| ||||| |||| ||  | || || ||| || ||||||| |
Db    1082  GCCCATTGGGGGATCTTCTTCATTCACTCCCAGAGCACATCCCAGGCTATTAAGCTGTT  1141

Qy    1704  TCGAGAGAAATTGCGCCTCTGAAAGAGAACAGTGAGCCACGTCAATGACCTTGCTCGCCCA  1763
                  |||||| || || | || |||| |||||||||||| || | || |||||| ||||| | |
Db    1142  CAAAGAAGAATTCTCCCCCCATGAAAGATGGAGTAAAGTTGGTGAATGATCTCGGCCCACCA  1201

Qy    1764  GCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAA  1823
                  || |||  ||  |||||| | ||  ||  |||| | ||  || |||| |||  | || |||
Db    1202  ACTTGCCATTTCTGTATGCTCACTTGTCAATGGCAATTCCCAGCCCTGGAACAGATCAA  1261

Qy    1824  CACCAGATGGAAGCTTCTGCAGGTGGCGCGTCGAGGACCGAGCAGCTGCATGAAGC  1883
                  |||| || |||| |||  |   |||||| |||  ||  ||  ||| || ||| ||||
Db    1262  CGTCCGATGAAACAACTACAGGCGCTCAGTTGATGAGAGCAGCTTAAGCAGCTCCAGGATGC  1321

Qy    1884  CCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCGTCGTCGTCCCTG  1943
                  || |||||||| || |||||||||  ||||| |  |||||||||  ||| ||||||||
Db    1322  CCACCGGGACTTTGGGCCTGGGTCACAGCAGCACTTTCTCTCCTCGTCGTCCAGTTCCCTG  1381

Qy    1944  GGAGGAGCCATCTCGCCAAACAAACAAGTGCCCACTACTATCAACCACCAGGAGACTCAAAC  2000
                  |||||||  || ||  ||  |||||||||| ||  ||  | |||| ||| |||| ||  ||
Db    1382  GGAAAGAGCAATTTCACCCAATAAAGTTCCCTACTACACATCAACCAGGCTCAGAC  1438
```

RESULT 8
US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
; US-09-091-501B-5

Query Match   4.0%;   Score 79.4; DB 4; Length 200;
Best Local Similarity   64.3%; Pred. No. 8.8e-15;
Matches 119; Conservative   0; Mismatches 66; Indels   0; Gaps   0;

```
Qy     465  CCTAAAACGCCAAGTACACACATAGGTGTCTTCAAGAAGATCAGAACAAGAACAAGT  524
                  || ||| ||||||  ||||| ||||||||| |||| ||||||||| |||||||||
Db      16  CCTGCAAAACCTGCTTGAAGAACATAAAAAGTTGCAAAGTGACCTGAAGCTGAGCAGGT  75

Qy     525  CAGGGTCAATTCTCTCCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGC  584
                  |||||| | ||| |||| |||  |||||| || |||  ||| ||||| || || ||| |
Db      76  GAAGGTGAATTCCTTAACTCATATGGTGGTGTGATTGTGTGGGATGAAAAACAGTGGGAGAGCGC  135

Qy     585  AACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGCAAACATCTGTAG  644
                  |||||||||  ||| | || | ||| | |   ||||||| || | | | |||||| |||
Db     136  CACAGCTGTTTGGAAGATCAGTTACAGAGAAACTGGGTGAGCGCTGGACCAGCTGTATGCCG  195

Qy     645  ATGGA  649
                  |||||
Db     196  CTGGA  200
```

RESULT 9
US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
; US-09-091-501B-4

Query Match   3.9%;   Score 78.6; DB 4; Length 200;
Best Local Similarity   62.4%; Pred. No. 1.6e-14;
Matches 123; Conservative   0; Mismatches 74; Indels   0; Gaps   0;

```
Qy     453  TGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCAAGAAGATCTAGA  512
                  || |||| | ||  ||||| ||||||| |||||| || || |||||||| ||| |
Db       4  TGACCTGCCTCCCTGCGAGAAGCTGCTTCAAGAACAACATAAAAGTTTGCAAAATGACCTTGA  63

Qy     513  ACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGTTGATGAATCTAG  572
                  ||| | ||| || | ||| | |||| || ||| || |||||||| |||| | |||||||
Db      64  AGCTGAACAGGTGAAGGTAAAATCCTTAACTCCTTAACTCACCATGGTGGTGTGATTGGGATGATGAAAAACAG  123

Qy     573  TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC  632
                  ||| | ||| |||||||  | |||||||| |||| || | |||||||||| || |||
Db     124  TGGGGAGAGTGCCACAGCTCTTCTTCGGAAGATCAGTCAGTTACAGAAACTGGGTGAGCGCTGGAC  183

Qy     633  AAACATCTGTAGATGGA  649
                  ||| |||| |||||||
Db     184  AGCTGTATGCCGCTGGA  200
```

RESULT 10
US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression

; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match   3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity   62.4%; Pred. No. 1.6e-14;
Matches 123; Conservative   0; Mismatches 74; Indels   0; Gaps   0;

Qy   453 TGATCTTGAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCAAGAAGATCTAGA 512
          ||||||||| ||||||||| || ||| || |||||||||
Db   4 TGATGTGAAATCTCTACAAAAGCTGCTAGAAGAACATAAAAGTTTGCAAAGTGATCTTGA 63

Qy   513 ACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGTTGATGAATCTAG 572
          |||||||||||| |||| |||||||| |||||||
Db   64 GGCTGAACAGGTGAAAGTAAATTCATTCACTAACTCACATGTGGTCATTGTTGATGAAAACAG 123

Qy   573 TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGGC 632
          ||||| |||
Db   124 TGGTGAGAGCGCTACAGCTTATCCTAGAAGAACCAGTTACAGAAACTTGGTGAGCGCTGGAC 183

Qy   633 AAACATCTGTAGATGGA 649
          || ||
Db   184 AGCAGTATGCCGTTGGA 200

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match   3.8%; Score 76.6; DB 1; Length 7218;
Best Local Similarity   6.7%; Pred. No. 8.8e-13;
Matches   28; Conservative 236; Mismatches 155; Indels   0; Gaps   0;

Qy   113 TTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGG 172
          |||| | | | || |||| || ||| || :: ||| :::::: |: RRRRRRRRR
Db   1474 TATCTATGCAAGTAGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRR 1415

Qy   173 ATTTGACAGCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTG 232
          : ||||||RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Db   1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355

Qy   233 GAACAGGAAAATTATCAGAAAGTGAAGAAACTGAAGTACAGAGAGCAGAGATGAATCTCCTAA 292
          :: :::::::RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Db   1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295

Qy   293 ATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTACATAGAG 352
          ::: :::::::RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Db   1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235

Qy   353 TTTTAATGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAG 412
          ::: ::::RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Db   1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175

Qy   413 AAGAAAGAACAACAGGAAAATGAGGAGGAGCAGCCTCTTGGACCTGATCTTGAACCTAAAAC 472
          ::::::::::: :::::::::::   ::: ::
Db   1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

Qy   473 GCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTC 531
          :::: :: ::::::
Db   1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTC 1056

RESULT 12
US-09-687-875A-13
; Sequence 13, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED P
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pXX-C2 5' junction
US-09-687-875A-13

Query Match   3.2%; Score 63.6; DB 4; Length 238;

Best Local Similarity 94.3%; Pred. No. 8.8e-10;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1065 TTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGA 1124
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   169 TTCGGACGACCAGTACTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGGAGTGA 228

Qy 1125 ACTTCGGAAA 1134
         ||||||||||
Db   229 ACTTCGGAAA 238

RESULT 13
US-09-621-976-15639/c
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match   2.3%; Score 46.2; DB 4; Length 505;
Best Local Similarity 17.4%; Pred. No. 0.00042;
Matches 54; Conservative 131; Mismatches 124; Indels 2; Gaps 1;

Qy   219 AAGTAAGCTGATTGGAACAGCAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCA 278
         || |:|: : | | ::| | ||:| | | ::|:|:|:|:|
Db   445 AATYAACRWAWANGAGAAWKRWKMAGSMRACARAGTTMAWGACWAMARRGWRCARGR--T 388

Qy   279 GATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGCTAGCATGGAAAAAACAAAG 338
         | :|:| ||||: || :::||:|:|| || ||:: ||:| |:|| :|||:| |::|
Db   387 GSWGSWKGGYRMWGRGMAAAKMRRMAAAGGSYCGMTSYTSGSKMTGRKGSMTKRKRMMTY 328

Qy   339 CAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTG 398
         : :::|::| | |:::||:|: | : ||:::: |: |: ::::::
Db   327 SGMWWTSYKCTKTGKKYTGWKSKKTRWTCTSWRKYMWNMSGCWARSMKSWARSWYSMMAC 268

Qy   399 GCTAACAAAAACAGAAGAAGAACAAGGAAAAATGGAGGAAGAGCCCTCTTGGACCTGATCT 458
         |:::||:||| | :|| | |:|:|| |:|| ::| ::::|
Db   267 WCMMSASAYRARRSMYGARRSMMRAGAGWWRAARGKKRARGKSSMMRSKMSSMRMSAGKA 208

Qy   459 TGAAGACCTAAAACGCCAAGTACACAACATAAGGTGCTTCAGAAGATCTAGAACAAGA 518
         :::|:: :::: || :| ::| :|| ||:|::::: ::|:
Db   207 RMCRRWMWSCRRMSYSCMGSKCMSCRGTCAKMWRYARYAKRYASSMGKYMGCRWCYAKC 148

Qy   519 ACAAGTCAGGG 529
         |:| |:::::|
Db   147 ARMYGYYRSRS 137

RESULT 14
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24

; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature

```
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
```

```
Query Match           2.2%;      Score 44;   DB 4;   Length 1230025;
Best Local Similarity  53.5%;
Matches  92;  Conservative  0;  Mismatches 80; Indels  0;  Gaps  0;

Qy    242 AATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGAT 301
          ||||| |||||| |||||  |||| || ||||||||| ||| ||||| |||||| ||
Db 656477 AATTAGAAGAAAGAAGAAGAAGAAGAATTGAGGATATCAAAGACTCAGATACAAAAT 656418

Qy    302 GGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGG 361
          |||  ||  |||| | |||| || |||||| ||||| ||||| |||||||| |||||
Db 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCATCACTAGGCAAGCAATTTATGTGG 656358

Qy    362 ATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGA 413
          ||| || || |||| ||||  ||||||| |||| |||||   |||||||
Db 656357 CAATTAAGCAGAAAAAACTAAAGCTTCTAAGAGACGCCTGGGAAATAGA 656306
```

```
                STREET: 100 Beaver Street
                CITY: Waltham
                STATE: Massachusetts
                COUNTRY: USA
                ZIP: 02354
        COMPUTER READABLE FORM:
                MEDIUM TYPE: CD/ROM ISO9660
                COMPUTER: PC
                OPERATING SYSTEM: <Unknown>
                SOFTWARE: ASCII
        CURRENT APPLICATION DATA:
                APPLICATION NUMBER: US/09/107,532A
                FILING DATE: 30-Jun-1998
        PRIOR APPLICATION DATA:
                APPLICATION NUMBER: 60/085,598
                FILING DATE: 14 May 1998
                APPLICATION NUMBER: 60/051571
                FILING DATE: July 2, 1997
        ATTORNEY/AGENT INFORMATION:
                NAME: Ariniello, Pamela Deneke
                REGISTRATION NUMBER: 40,489
                REFERENCE/DOCKET NUMBER: GTC-012
        TELECOMMUNICATION INFORMATION:
                TELEPHONE: (781)893-5007
                TELEFAX: (781)893-8277
        INFORMATION FOR SEQ ID NO: 1186:
        SEQUENCE CHARACTERISTICS:
                LENGTH: 1179 base pairs
                TYPE: nucleic acid
                STRANDEDNESS: double
                TOPOLOGY: circular
        MOLECULE TYPE: DNA (genomic)
        HYPOTHETICAL: NO
        ANTI-SENSE: NO
        ORIGINAL SOURCE:
                ORGANISM: Enterococcus faecium
        FEATURE:
                NAME/KEY: misc_feature
                LOCATION: (B) LOCATION 1...1179
        SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186

Query Match        2.2%;  Score 43.4;  DB 4;  Length 1179;
Best Local Similarity   50.7%;  Pred. No. 0.0058;
Matches  104;  Conservative    0;  Mismatches  101;  Indels    0;  Gaps    0;

Qy    371 ATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAA 430
              |  ||||    ||   ||  |   |||| | |   |     ||   |    ||||
Db    602 AGCAAAGAATAAAAGATTTGATCGATCAGACACAAAGAAAAATGGAGATACGATCGGAGGAA 661

Qy    431 TGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATA 490
              ||| |||  |||   |||  ||  | |  ||   ||||   ||   |||| ||||||
Db    662 TTGTAGAAGTGCTCGTTGGAGGCGGTTCCAGCTGGATTAGGAAGCTACGTACAATGGGACA 721

Qy    491 AGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGG 550
              ||| ||| |||| ||| |  |   || |||  || ||  |  || ||  ||  |||
Db    722 CGAAGCTAGATGCCAAAATCGCACAAGCTGTGGTTAGTATCAATGCCTTTAAAGGCGTAG 781

Qy    551 TGGTGGTAGTTGATGAATCTAGTGG 575
              || ||  || || | |||  ||
Db    782 AATTTGGGGTCGGATTCACTTCTGG 806

Search completed: April  5, 2004, 08:24:49
Job time : 100.041 secs
```

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 04:53:30 ; Search time 485.207 Seconds
(without alignments)
15454.047 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.........tcaaccacgagactcaaaca 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2470430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters: 4940860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	DB	Match Length	ID	Description	
1	2001	100.0	10	3999	US-09-845-416-6	Sequence 6, Appli
2	2001	100.0	10	4966	US-09-845-416-28	Sequence 28, Appl
3	2001	100.0	10	4990	US-09-845-416-34	Sequence 34, Appl
4	1709	85.4	10	3858	US-09-845-416-9	Sequence 9, Appli
5	1709	85.4	10	4825	US-09-845-416-29	Sequence 29, Appl
6	1709	85.4	10	4848	US-09-845-416-35	Sequence 35, Appl
7	1709	85.4	10	5060	US-09-845-416-36	Sequence 36, Appl
8	1662.8	83.1	10	4182	US-09-845-416-2	Sequence 2, Appli
9	1662.8	83.1	10	5149	US-09-845-416-27	Sequence 27, Appl
10	1283	64.1	10	5462	US-09-845-416-41	Sequence 41, Appl
11	1182.6	59.1	15	8689	US-10-149-736-42	Sequence 42, Appl
12	1182.6	59.1	10	11058	US-09-845-416-1	Sequence 1, Appl
13	1182.6	59.1	15	11443	US-10-149-736-44	Sequence 44, Appl
14	1182.6	59.1	15	12057	US-10-149-736-47	Sequence 47, Appl
15	1182.6	59.1	9	13957	US-09-782-378A-22	Sequence 22, Appl

16	59.1	1182.6	13957	9	US-09-880-107-2284	Sequence 2284, Ap
17	59.1	1182.6	13957	15	US-10-149-736-1	Sequence 1, Appli
18	59.1	1182.6	14069	12	US-10-342-887-434	Sequence 434, App
19	59.1	1182.6	14082	12	US-10-342-887-981	Sequence 981, App
20	59.1	1182.6	14082	15	US-10-341-434-108	Sequence 108, App
21	59.1	1182	2169	15	US-09-845-416-4	Sequence 4, Appli
22	59.1	1182	3531	10	US-09-845-416-10	Sequence 10, Appl
23	59.1	1182	4498	10	US-09-845-416-30	Sequence 30, Appl
24	59.0	1180	5339	10	US-10-149-736-40	Sequence 40, Appl
25	50.6	1013	3510	15	US-09-845-416-12	Sequence 12, Appl
26	50.6	1013	4476	10	US-09-845-416-31	Sequence 31, Appl
27	50.2	1004	13815	15	US-10-149-736-2	Sequence 2, Appl
28	41.7	835	1821	10	US-09-845-416-13	Sequence 13, Appl
29	36.3	727	5417	15	US-10-149-736-39	Sequence 39, Appl
30	33.9	678.4	4414	15	US-10-342-887-416-32	Sequence 32, Appl
31	33.9	678	1991	10	US-09-845-416-3	Sequence 3, Appli
32	33.8	677	1667	10	US-09-845-416-7	Sequence 7, Appli
33	33.3	666.4	3446	10	US-09-845-416-14	Sequence 14, Appl
34	22.4	448	1434	10	US-09-845-416-15	Sequence 15, Appl
35	20.4	408.2	11096	15	US-10-149-736-4	Sequence 4, Appli
36	19.8	397	10302	9	US-09-782-378A-23	Sequence 23, Appl
37	19.8	397	10302	15	US-10-149-736-3	Sequence 3, Appl
38	19.3	387	387	15	US-10-149-736-32	Sequence 32, Appl
39	19.2	385	16531	14	US-10-101-510-667	Sequence 667, App
40	17.5	350	1340	10	US-09-845-416-11	Sequence 11, Appl
41	17.4	348	348	15	US-10-149-736-31	Sequence 31, Appl
42	16.5	330	333	15	US-10-149-736-9	Sequence 9, Appl
43	16.3	327	327	15	US-10-149-736-8	Sequence 8, Appl
44	16.2	324	324	15	US-10-149-736-33	Sequence 33, Appl
45	16.1	322.6	5106	12	US-10-220-120-157	Sequence 157, App

ALIGNMENTS

RESULT 1
US-09-845-416-6
; Sequence 6, Application US/0984541
6
; Publication No. US2003017131
2A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6

Query Match   100.0%;   Score 2001; DB 10; Length 3999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GGCAGTTCATTGATGGAGAGTGAGTAAACCTGACCGGTTATCAAACAGCTTTAGAGAA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1000 GGCAGTTCATTGATGGAGAGTGAGTAAACCTGACCGGTTATCAAACAGCTTTAGAGAA 1059

Qy   61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1119

Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGTACATGAGGTACATGAGATTGGACA 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1120 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGTACATGAGGTACATGAGATTGGACA 1179

Qy   181 GCCCATCAGGGCCGGGTTGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240

Right column:

```
Db  2260 ATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAGAACGATGTACATAGGGCCTTCAGAGG 2319
Qy  1321 GAATTGAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTGACGAATATTTCTG 1380
Db  2320 GAATTGAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTGACGAATATTTCTG 2379
Qy  1381 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCT 1440
Db  2380 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCT 2439
Qy  1441 GAGGAGGAGCCCAGAAGTCACTCGGCTTCTACGAAAGCAGCAGGTCGAGGAGGTCAATACT 1500
Db  2440 GAGGAGGAGCCCAGAAGTCACTCGGCTTCTACGAAAGCAGCAGGTCGAGGAGGTCAATACT 2499
Qy  1501 GAGTGGGAAAAATTGAACCTGACTCCGCCACTCCGTGACTGGCAGAGAAAATAGATGAGACCCTT 1560
Db  2500 GAGTGGGAAAAATTGAACCTGACTCCGCCACTCCGTGACTGGCAGAGAAAATAGATGAGACCCTT 2559
Qy  1561 GAAAGACTCCAGGACTTCAAGAGCCACGGATGAGCTGGACCTGGACCTCCTCCTCCAAGAT 1620
Db  2560 GAAAGACTCCAGGACTTCAAGAGCCACGGATGAGCTGGACCTGGACCTCCTCCTCCAAGAT 2619
Qy  1621 GAGGTGATCAAGGAGTCCTGGCAGCCCGTGGGGCCATCTCCATTGACTCTCTCCAAGAT 1680
Db  2620 GAGGTGATCAAGGAGTCCTGGCAGCCCGTGGGGCCATCTCCATTGACTCTCTCCAAGAT 2679
Qy  1681 CACCTTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGC 1740
Db  2680 CACCTTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGC 2739
Qy  1741 CACGTCAATGACCTTGCTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 1800
Db  2740 CACGTCAATGACCTTGCTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 2799
Qy  1801 CTCAGGACTCTGGAAGACTTGAACACCAGATGAGGAAGCTTCGCAGGTGGCCGTGGAGGAC 1860
Db  2800 CTCAGGACTCTGGAAGACTTGAACACCAGATGAGGAAGCTTCGCAGGTGGCCGTGGAGGAC 2859
Qy  1861 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCTCAGCACTTTCTT 1920
Db  2860 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCTCAGCACTTTCTT 2919
Qy  1921 TCCAGGTCGTCCAGGTCCTGGGAGAGACCCATCTGCCAAACAAAGTGCCCTACTAT 1980
Db  2920 TCCAGGTCGTCCAGGTCCTGGGAGAGACCCATCTGCCCAAACAAAGTGCCCTACTAT 2979
Qy  1981 ATCAACCACGAGACTCAAACA 2001
Db  2980 ATCAACCACGAGACTCAAACA 3000
```

Query Match 100.0%; Score 2001; DB 10; Length 4966;

Left column:

```
Db  1180 GCCCATCAGGGCCGGGTTGGTAATATTCACAATTGGGAAGTAAGCTGATTGGAACAGGA 1239
Qy   241 AAATTATCAGAGATGAAGAAACTGAAGTACAGAGCAGTGAATCCTAAATTCAAGA 300
Db  1240 AAATTATCAGAGATGAAGAAACTGAAGTACAGAGCAGTGAATCCTAAATTCAAGA 1299
Qy   301 TGGGAATGCCTCAGGTAGCTAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATG 360
Db  1300 TGGGAATGCCTCAGGTAGCTAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATG 1359
Qy   361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
Db  1360 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 1419
Qy   421 ACAAGGAAAATGGAGGAAGAGCCCTTGAACCTGATCTTGAAGACCTAAAACGGCCAGTA 480
Db  1420 ACAAGGAAAATGGAGGAAGAGCCCTTGAACCTGATCTTGAAGACCTAAAACGGCCAGTA 1479
Qy   481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAGAACAAGTCAGGGTCAATTCTTC 540
Db  1480 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAGAACAAGTCAGGGTCAATTCTTC 1539
Qy   541 ACTCACATGGTGGTGTGTAGTTGATGAATCAGTGGAGATCACGCAACTGTCTGCTTTGGAA 600
Db  1540 ACTCACATGGTGGTGTGTAGTTGATGAATCAGTGGAGATCACGCAACTGTCTGCTTTGGAA 1599
Qy   601 GAACAACTTAAGGTATTGGGAGATCGATGGCAAACATCTGTAGATGGACAGAAGACCGC 660
Db  1600 GAACAACTTAAGGTATTGGGAGATCGATGGCAAACATCTGTAGATGGACAGAAGACCGC 1659
Qy   661 TGGGTTCTTTTACAAGACCACCGCCTAGCTCCTGGACTGACCATTGGAGGCCTCT 720
Db  1660 TGGGTTCTTTTACAAGACCACCGCCTAGCTCCTGGACTGACCATTGGAGGCCTCT 1719
Qy   721 CCTACTCCAGACTGTTACTCTGGTGACAACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 780
Db  1720 CCTACTCCAGACTGTTACTCTGGTGACAACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 1779
Qy   781 AAACTAGAAAATGCCATCTCCTTCCTTGATGTTGGAGTGGTTACCTCATAGATTACTGCAACAG 840
Db  1780 AAACTAGAAAATGCCATCTCCTTCCTTGATGTTGGAGTGGTTACCTCATAGATTACTGCAACAG 1839
Qy   841 TTCCCCCTGGAACTGGAAAAGTTTCTTCTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900
Db  1840 TTCCCCCTGGAACTGGAAAAGTTTCTTCTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 1899
Qy   901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAG 960
Db  1900 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAG 1959
Qy   961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 1020
Db  1960 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 2019
Qy  1021 AACCTGGATGAAAACAGCCAAAAATCCTGAGATCCTGAGATCCTGAAGGTTCCGATGATGCAGTC 1080
Db  2020 AACCTGGATGAAAACAGCCAAAAATCCTGAGATCCTGAGATCCTGAAGGTTCCGATGATGCAGTC 2079
Qy  1081 CTGTTACAAAGACGTTTGGATAACTAACTTCAAGTGAGTGGAGTGAACTTCGGAAAAAGTCT 1140
Db  2080 CTGTTACAAAGACGTTTGGATAACTAACTTCAAGTGAGTGGAGTGAACTTCGGAAAAAGTCT 2139
Qy  1141 CTCAACATTAGGTGTCCATTGGAAGCCAGTTGGACCAGTGGAAGCGTCTGCACTTTCT 1200
Db  2140 CTCAACATTAGGTGTCCATTGGAAGCCAGTTGGACCAGTGGAAGCGTCTGCACTTTTCT 2199
Qy  1201 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCCACCT 1260
Db  2200 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCCACCT 2259
Qy  1261 ATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAGAACGATGTACATAGGGCCTTCAGAGG 1320
```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GGCAGTTCATTGATGGAGGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1757 GGCAGTTCATTGATGGAGGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAA 1816

Qy     61 GTATTATCGTGGCTTCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1817 GTATTATCGTGGCTTCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1876

Qy    121 GATGTGGAAGTGGTGAAAGACCAGTTCATCATCTCATGAGGGGTACACTGATGGATTTGACA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1877 GATGTGGAAGTGGTGAAAGACCAGTTCATCATCTCATGAGGGGTACACTGATGGATTTGACA 1936

Qy    181 GCCCATCAGGGCGCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1937 GCCCATCAGGGCGCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996

Qy    241 AAATTATCAGAAGAGATGAAGAAACTGAAGTACAAGAGACGAGATGAATCTCCTAAATTCAAGA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1997 AAATTATCAGAAGAGATGAAGAAACTGAAGTACAAGAGACGAGATGAATCTCCTAAATTCAAGA 2056

Qy    301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 2116

Qy    361 GATCTCCAGAATCAGAAATCGAAAGAGTTGAATGACTCGGCTCAACAAAAACAGAAGAAAGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2117 GATCTCCAGAATCAGAAATCGAAAGAGTTGAATGACTCGGCTCAACAAAAACAGAAGAAAGA 2176

Qy    421 ACAAGGAAAATGGAGGAAGAGCCTCTTGAAGACCTAAAACGCCAAGTA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2177 ACAAGGAAAATGGAGGAAGAGCCTCTTGAAGACCTAAAACGCCAAGTA 2236

Qy    481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAACAAGAAGTCAGGGTCAATTCTCTC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2237 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAACAAGAAGTCAGGGTCAATTCTCTC 2296

Qy    541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2297 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAA 2356

Qy    601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGCAGCAGAGACCGC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2357 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGCAGCAGAGACCGC 2416

Qy    661 TGGGTTCTTTTACAAGACTGTTACTCTGGTGACCAACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2417 TGGGTTCTTTTACAAGACTGTTACTCTGGTGACCAACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 2476

Qy    721 CCTACTCAGACTGTTACTCTGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2477 CCTACTCAGACTGTTACTCTGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCC 2536

Qy    781 AAACTAGAAATGCCATCTCCTTGATGTTGAGGTACCTACTCATAGATTACTGCAACAG 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2537 AAACTAGAAATGCCATCTCCTTGATGTTGAGGTACCTACTCATAGATTACTGCAACAG 2596

Qy    841 TTCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTGCCTGGCTTACGAGAACTCCAGGG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2597 TTCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTGCCTGGCTTACGAGAACTCCAGGG 2656

Qy    901 GTCCTACGGATGCCATCTCCGTACCCGTAAGGAAAGGCTCCAGGAAGACTCCAAACA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2657 GTCCTACAGATGCCATCTCCGTACCCGTAAGGAAAGGCTCCAGGAAGACTCCAAACA 2716

Qy    961 CTGATGAAACAATGGCAAGACTTCCAGGTGAAATTGAAGCTGAAATTGAAGCTCACACAGATGTTTATCAC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2717 CTGATGAAACAATGGCAAGACTTCCAGGTGAAATTGAAGCTGAAATTGAAGCTCACACAGATGTTTATCAC 2776

Qy   1021 AACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2777 AACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTC 2836

Qy   1081 CTGTTACAAAGACGTTTGGATAACATGACTTCAAGTGGAGTGGAGTGAACTTCGGAAAAAGTCT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2837 CTGTTACAAAGACGTTTGGATAACATGACTTCAAGTGGAGTGGAGTGAACTTCGGAAAAAGTCT 2896

Qy   1141 CTCAACATTAGGTCCCATTTGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2897 CTCAACATTAGGTCCCATTTGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCT 2956

Qy   1201 CTGCAGGAACTTCTGTGTGTGGCTACAGCTGAAAGATGATGATGAATTAAGCCGGCAGCGACACCT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2957 CTGCAGGAACTTCTGTGTGTGGCTACAGCTGAAAGATGATGATGAATTAAGCCGGCAGCGACACCT 3016

Qy   1261 ATTGGAGGCGACTTTCAGCAGTTCGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3017 ATTGGAGGCGACTTTCAGCAGTTCGAAAGCAGAACGATGTACATAGGGCCTTCAAGAGG 3076

Qy   1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTTGTACGAATATTTCTG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3077 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTTGTACGAATATTTCTG 3136

Qy   1381 ACAGAGCAGCCTTGGAAGGACTAGAGAAACTCTACCAGGAGAGCCCAGAGAGCTGCCTCCT 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3137 ACAGAGCAGCCTTGGAAGGACTAGAGAAACTCTACCAGGAGAGCCCAGAGAGCTGCCTCCT 3196

Qy   1441 GAGGAGGAGCCCAGAATGTCACTCGGCTTCACGGAAAGCAGGCTGAGGAGGTCAATACT 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3197 GAGGAGGAGCCCAGAATGTCACTCGGCTTCACGGAAAGCAGGCTGAGGAGGTCAATACT 3256

Qy   1501 GAGTGGGGAAAATTGAACCTGACTCCGGTGACTGGCAGAGAAAAATAGATGAGACCCTT 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3257 GAGTGGGGAAAATTGAACCTGACTCCGGTGACTGGCAGAGAAAAATAGATGAGACCCTT 3316

Qy   1561 GAAAGACTCCAGGAACTTCAGGAGCCACGGATGAGCGTCGACCTCAAGTGCGCGCCAAGCT 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3317 GAAAGACTCCAGGAACTTCAGGAGCCACGGATGAGCGTCGACCTCAAGTGCGCGCCAAGCT 3376

Qy   1621 GAGGTGATCAAGGGATCCTGGCAGCCCCGGTGGGCGATCTCCTCATTGACTCTTCCAAGAT 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3377 GAGGTGATCAAGGGATCCTGGCAGCCCCGGTGGGCGATCTCCTCATTGACTCTTCCAAGAT 3436

Qy   1681 CACCTCGAAAGTCAAGGCACTTCGAGGAGAACGTGAGC 1740
            ||||||||||||||||||||||||||||||||||||||
Db   3437 CACCTCGAAAGTCAAGGCACTTCGAGGAGAACGTGAGC 3496

Qy   1741 CACGTCAATGACCTTGCTCGCGAGCTACACTTTGGGCATTCAGCTCTCACCGTATAAC 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3497 CACGTCAATGACCTTGCTCGCGAGCTACACTTTGGGCATTCAGCTCTCACCGTATAAC 3556

Qy   1801 CTCAGCACTCTGGAAGACCTGAAGCTTCTGCAGGTGGCGTCGAGGAC 1860
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   3557 CTCAGCACTCTGGAAGACCTGAAGCTTCTGCAGGTGGCGTCGAGGAC 3616

Qy   1861 CGGGTCAGGCAGCTGCATGAGCGTCCTCCAGGACTTTGGTCCAGCATCCAGCACTTTCTT 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3617 CGGGTCAGGCAGCTGCATGAGCGTCCTCCAGGACTTTGGTCCAGCATCCAGCACTTTCTT 3676

Qy   1921 TCCACGTCGTCCAGGGTCCCTGGGGAGGAGAGCCATCCGCCAAACAAAGTGCCCTACTAT 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3677 TCCACGTCGTCCAGGGTCCCTGGGGAGGAGAGCCATCCGCCAAACAAAGTGCCCTACTAT 3736

Qy   1981 ATCAACCACGAGACTCAAACA 2001
            |||||||||||||||||||||
Db   3737 ATCAACCACGAGACTCAAACA 3757
```

RESULT 3
US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO

```
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-34

Query Match    100.0%;    Score 2001;    DB 10;    Length 4990;
Best Local Similarity  100.0%;    Pred. No.0;
Matches 2001;    Conservative  0;    Mismatches  0;    Indels  0;    Gaps  0;
```

Qy   1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAA   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1781 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAA   1840

Qy   61 GTATTATCGTTGCCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAAT   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1841 GTATTATCGTTGCCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAAT   1900

Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTCATATCTGTGGCTACATGAGGGGTACATGATTGACA   180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1901 GATGTGGAAGTGGTGAAAGACCAGTTCATATCTGTGGCTACATGAGGGGTACATGATTGACA   1960

Qy   181 GCCCATCAGGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1961 GCCCATCAGGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA   2020

Qy   241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACCAGATGAATCTCCTAAATTCAAGA   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2021 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGACCAGATGAATCTCCTAAATTCAAGA   2080

Qy   301 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATG   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2081 TGGGAATGCCTCAGGGTAGCTAGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATG   2140

Qy   361 GATCTCCAGAATCAGAAACTGAATGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2141 GATCTCCAGAATCAGAAACTGAATGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA   2200

Qy   421 ACAAGGAAAATGGAGGAAGAGCCTCTTGAAGATGACTTGAGAACGCCAAGTA   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2201 ACAAGGAAAATGGAGGAAGAGCCTCTTGAAGATGACTTGAGAACGCCAAGTA   2260

Qy   481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2261 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC   2320

Qy   541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAAGATCACGCAACTGCTTGGAA   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2321 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAAGATCACGCAACTGCTTGGAA   2380

Qy   601 GAACAACTTAAGGTATTGGGAGATGCGGCAAACATCTGTAGATGGACAGAAGACCGC   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2381 GAACAACTTAAGGTATTGGGAGATGCGGCAAACATCTGTAGATGGACAGAAGACCGC   2440

Qy   661 TGGGTTCTTTACAAGACAGCCAGCCTGACCTAGCTCCTGACTGACCACTATTGGAGCCTCT   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2441 TGGGTTCTTTACAAGACAGCCAGCCTGACCTAGCTCCTGACTGACCACTATTGGAGCCTCT   2500

Qy   721 CCTACTCAGACTGTTACTTCTGTTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC   780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2501 CCTACTCAGACTGTTACTTCTGTTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC   2560

Qy   781 AAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2561 AAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG   2620

Qy   841 TTCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAGAAGCTGAAACAACTGCCAAT   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2621 TTCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAGAAGCTGAAACAACTGCCAAT   2680

Qy   901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAGACTCCAAGGGAGTAAAAGAG   960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2681 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTAGAAGACTCCAAGGGAGTAAAAGAG   2740

Qy   961 CTGATGAAACAATGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAC   1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741 CTGATGAAACAATGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAC   2800

Qy   1021 AACCTGGATGAAAACAGCCAAAAATCCTGAGATCCCTGGAAGTGGAGAAAAAGTCT   1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 AACCTGGATGAAAACAGCCAAAAATCCTGAGATCCCTGGAAGTGGAGAAAAAGTCT   2860

Qy   1081 CTGTTACAAAGACGTTGGATAACATGAACATGACTTCAAGTGGAGTGAACTTCGAAAAAGTCT   1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 CTGTTACAAAGACGTTGGATAACATGAACATGACTTCAAGTGGAGTGAACTTCGAAAAAGTCT   2920

Qy   1141 CTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGCCGTCTGACCCTTTCT   1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 CTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGCCGTCTGACCCTTTCT   2980

Qy   1201 CTGCAGGAACTTCTGTGTGCTACAGCTGAAAGATGATGAAAGATGATGAATTGACA   1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 CTGCAGGAACTTCTGTGTGCTACAGCTGAAAGATGATGAAAGATGATGAATTGACA   3040

Qy   1261 ATTGGAGGCGGACTTTCCAGCAGTTCAGAAGCAGAACAGATGTACATAGGGCCTTCAAGAGG   1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 ATTGGAGGCGGACTTTCCAGCAGTTCAGAAGCAGAACAGATGTACATAGGGCCTTCAAGAGG   3100

Qy   1321 GAATTGAAAACTAAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTG   1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 GAATTGAAAACTAAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTG   3160

Qy   1381 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTCACCAGGAGCCCAGAGAGTGCCTCCT   1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTCACCAGGAGCCCAGAGAGTGCCTCCT   3220

Qy   1441 GAGGAGAGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGTCAATACT   1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GAGGAGAGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGTCAATACT   3280

Qy   1501 GAGTGGGAAAAATTGAACCTGACTGGCACTCCGCTGCACTCCGCTGCACTTGGCAGAGAAAAATAGATGAGACCCTT   1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 GAGTGGGAAAAATTGAACCTGACTGGCACTCCGCTGCACTCCGCTGCACTTGGCAGAGAAAAATAGATGAGACCCTT   3340

Qy   1561 GAAAGACTCCAGGGACTTCAAGAGAGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT   1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 GAAAGACTCCAGGGACTTCAAGAGAGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT   3400

Qy   1621 GAGGTGATCAAGGGATCCTGGGAGATCTCTCCAGAT   1680
            ||||||||||||||||||||||||||||||||||||
Db   3401 GAGGTGATCAAGGGATCCTGGGAGATCTCTCCAGAT   3460

Qy   1681 CACCTGAGAAGTCAAGGCACTTCCGCCTTCGAGGAGAAATTGCGCCTCTCTGAAAGAACGTGAGC   1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 CACCTGAGAAGTCAAGGCACTTCCGCCTTCGAGGAGAAATTGCGCCTCTCTGAAAGAACGTGAGC   3520

Qy   1741 CACGTCAATGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC   1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 CACGTCAATGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC   3580

Qy   1801 CTCAGCACTCTGAAGACCTGAACACCAGATGAACTGCCATCTCC   1860
            ||||||||||||||||||||||||||||||||||||||||||||
Db   3581 CTCAGCACTCTGAAGACCTGAACACCAGATGAACTGCCATCTCC   3640

Qy   1861 CGAGTCAGGCAGCTGCAGTCTGATGAAGCCCACAGGGACTTTGGTCAGTCCAGCACTTTCTT   1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 CGAGTCAGGCAGCTGCAGTCTGATGAAGCCCACAGGGACTTTGGTCAGTCCAGCACTTTCTT   3700

RESULT 4

US-09-845-416-9
Sequence 9, Application US/0984416
Publication No. US2003017131211
GENERAL INFORMATION:
  APPLICANT: XIAO, XIAO
  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
  TITLE OF INVENTION: THEREOF
  FILE REFERENCE: DE1142
  CURRENT APPLICATION NUMBER: US/09/845,416
  CURRENT FILING DATE: 2001-04-30
  PRIOR APPLICATION NUMBER: 60/200,777
  PRIOR FILING DATE: 2000-04-28
  NUMBER OF SEQ ID NOS: 36
  SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9
  LENGTH: 3858
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-845-416-9

Query Match   85.4%;   Score 1709;   DB 10;   Length 3858;
Best Local Similarity   93.0%;   Pred. No. 0;
Matches 1860;   Mismatches   0;   Conservative   0;   Indels   141;   Gaps   1;

Qy   1 GGCAGTTCATTGATGAGGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTAGAAGA 60
            |||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1000 GGCAGTTCATTGATGAGGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTAGAAGA 1059

Qy   61 GTATTATCGTGGCTCTTCTTCTGCTGAGGACACATTGCAAGCACAGGGAGATTTCTAAT 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1060 GTATTATCGTGGCTCTTCTTCTGCTGAGGACACATTGCAAGCACAGGGAGATTTCTAAT 1119

Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGTACATGATGGATTGACA 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1120 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGTACATGATGGATTGACA 1179

Qy   181 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1180 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1239

Qy   241 AAATTATCAGAAGATGAAGAAAACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1240 AAATTATCAGAAGATGAAGAAAACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGA 1299

Qy   301 TGGGAATGCCTCAGGGTAGCTAGCATCGAAAAACAAAGCAATTTACATGAGTTTTAATG 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1300 TGGGAATGCCTCAGGGTAGCTAGCATCGAAAAACAAAGCAATTTACATGAGTTTTAATG 1359

Qy   361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1360 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 1419

Qy   421 ACAAGGAAAATGAGGAAGAGAGCCTCTTGACCTGATCTTGAAGAACCTAAAACGCCAAGTA 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1420 ACAAGGAAAATGAGGAAGAGAGCCTCTTGACCTGATCTTGAAGAACCTAAAACGCCAAGTA 1479

Qy   481 CAACAACATAAGGTGCTTCAAGACAGTCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1480 CAACAACATAAGGTGCTTCAAGACAGTCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 1539

Qy   541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTTGCTTGGAA 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTTGCTTGGAA 1599

Qy   1921 TCCACGTCGTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 TCCACGTCGTGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTAT 3760

Qy   1981 ATCAACCACGAGACTCAAACA 2001
            |||||||||||||||||||||
Db   3761 ATCAACCACGAGACTCAAACA 3781

Db   1540 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTTGCTTGGAA 1599

Qy   601 GAACAACTTAAGGTATTGGAGATCGATGGGCAACAAACATCTGTAGATGGACAGAAGACCGC 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1600 GAACAACTTAAGGTATTGGAGATCGATGGGCAACAAACATCTGTAGATGGACAGAAGACCGC 1659

Qy   661 TGGGTTCTTTACAAGACACCAGCCTAGCTCCTGGACTGACCACTATTGGAGGCCTCT 720
            |||||||||||||||||||||||||||||||
Db   1660 TGGGTTCTTTTACAAGAC------------- 1677

Qy   721 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCC 780
Db   1678 --- 1677

Qy   781 AAACTAGAATGCCATCTTCCTTGATGTTGGAGTGACCTACCTACTCATAGATTACTGCAACAG 840
                                            ---ACTCATAGATTACTGCAACAG 1698
Db   1678 --- 1698

Qy   841 TTCCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAAT 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 TTCCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAAT 1758

Qy   901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGAACACTCCAAGGGAGTAAAAGAG 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGAACACTCCAAGGGAGTAAAAGAG 1818

Qy   961 CTGATGAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAC 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 CTGATGAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCAC 1878

Qy   1021 AACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCTGAGAATCCTCAAGTGAGTTCCGATGATGCAGTC 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 AACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCTGAGAATCCTCAAGTGAGTTCCGATGATGCAGTC 1938

Qy   1081 CTGTTACAAAGACGTTTGGATAACAGCCAAGTGAGTGAACTTCAAGGAAAAAGTCT 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 CTGTTACAAAGACGTTTGGATAACAGCCAAGTGAGTGAACTTCAAGGAAAAAGTCT 1998

Qy   1141 CTCAACATTAGGTCCCCATTTGGAAGCCAGTTCGACCAGTGGAAGGGTCTCGCACCTTTCT 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 CTCAACATTAGGTCCCCATTTGGAAGCCAGTTCGACCAGTGGAAGGGTCTCGCACCTTTCT 2058

Qy   1201 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCCACCT 1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCCACCT 2118

Qy   1261 ATTGGAGGCGACTTTCCAGCAGTCGAGAAGCAGAAGCAGATGTACATAGGCGCTTCAAGAGG 1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 ATTGGAGGCGACTTTCCAGCAGTCGAGAAGCAGAAGCAGATGTACATAGGCGCTTCAAGAGG 2178

Qy   1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGACCTCTTGAGACTGTACTCTTGAGAATATTTCTG 1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2179 GAATTGAAAACTAAAGAACCTGTAATCATGAGACCTCTTGAGACTGTACTCTTGAGAATATTTCTG 2238

Qy   1381 ACAGAGCAGCCTTGGAAGGACTAGAGGAACTCTACCAGGAAGAGCCTGCCTCCT 1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2239 ACAGAGCAGCCTTGGAAGGACTAGAGGAACTCTACCAGGAAGAGCCTGCCTCCT 2298

Qy   1441 GAGGAGAGCCCAGAGTGTCACTCGGCTTCACGAAAGACAGGCTGAGGAGTCAATACT 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2299 GAGGAGAGCCCAGAATGTCACTCGGCTTCACGAAAGACAGGCTGAGGAGTCAATACT 2358

Qy   1501 GAGTGGAAAAATTGAACCTGACTCGCTGACTCGGCAGCAGAAAATAGATGAGACCCTT 1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2359 GAGTGGGAAAAATTGAACCTGACTCGCTGACTCGGCAGCAGAAAATAGATGAGACCCTT 2418

Qy   1561 GAAAGACTTCCAGAACTTCAGGAACTTCAAGAGCTGAACCTGACTCTCCATTGACTCCTCCAAGAT 1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2419 GAAAGACTTCCAGAACTTCAGGAACTTCAAGAGCTGAACCTGACTCTCCATTGACTCCTCCAAGCT 2478

Qy   1621 GAGGTGATCAAGGGATCCTGGCGCAGCCCGGGATCCTCCATTGACTCCTCCAAGAT 1680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2479 GAGGTGATCAAGGGATCCTGGCGCAGCCCGGGATCCTCCATTGACTCCTCCAAGAT 2538

```
Qy   1681 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2539 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGC 2598

Qy   1741 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCCACCGTATAAC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2599 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCCACCGTATAAC 2658

Qy   1801 CTCAGCACTCTGGAGAACCTGAACACCAGAGAGCTTCTGCAGGTGGCCGTCGAGGAC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2659 CTCAGCACTCTGGAGAACCTGAACACCAGAGAGCTTCTGCAGGTGGCCGTCGAGGAC 2718

Qy   1861 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTT 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2719 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCTCAGCACTTTCTT 2778

Qy   1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2779 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 2838

Qy   1981 ATCAACCACGAGACTCAAACA 2001
          |||||||||||||||||||||
Db   2839 ATCAACCACGAGACTCAAACA 2859
```

RESULT 5
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match   85.4%; Score 1709; DB 10; Length 4825;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

```
Qy      1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTAAACCGTTATCAAACAGCTTTAGAAGAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1757 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAAGAA 1816

Qy     61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1876

Qy    121 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGGATTTGACA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1877 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGGATTTGACA 1936

Qy    181 GCCCATCAGGCGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1937 GCCCATCAGGCGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 1996

Qy    241 AAATTATCAGAAAACTAAAGAACCTGTAATCATGAAGATCTCTAAATTCAAGA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1997 AAATTATCAGAAAACTAAAGAACCTGTAATCATGAAGATCTCCTAAATTCAAGA 2056

Qy    301 TGGGAATGCCTCAGGGTAGCTAGCATGGAGAAAAACAAAGCAATTTACATAGAGTTTTAATG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2996 ACAGAGCAGCGCTTTGGAAGGACTAGAGAACTTCACCAGGAGACTAGAGAACTTCACCAGGAGAGCTGCCTCCT 3055
```

```
Db   2057 TGGGAATGCCTCAGGGTAGCTAGCATGGAGAAAAACAAAGCAATTTACATAGAGTTTTAATG 2116

Qy    361 GATCTTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAAAGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2117 GATCTTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAAAGA 2176

Qy    421 ACAAGGAAAATGGAGGAGAGGAGGCCTCTTGGACCTGATCTTGAGAACGCCAAGTA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2177 ACAAGGAAAATGGAGGAGAGGAGGCCTCTTGGACCTGATCTTGAGAACGCCAAGTA 2236

Qy    481 CAACAACTAAGGTGCTTCAAGAGATCTAGAACAAGAGCTAGAACAAGTCAGGGTCAATTCTCTC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2237 CAACAACATAAGGTGCTTCAAGAGATCTAGAAGAAGATCTAGAACAAGTCAGGGTCAATTCTCTC 2296

Qy    541 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGGATCACGCAACTGCTTGCTTTGGAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2297 ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCAGGATCACGCAACTGCTTGCTTTGGAA 2356

Qy    601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAAGAGAACCGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2357 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAAGAGAACCGC 2416

Qy    661 TGGGTTCTTTACAAGACCAGCCTGACCTGGTGTTACTTAAGGAAACTGCCATCTCC 720
          |||||||||||||||||||||||||||||||||||||||
Db   2417 TGGGTTCTTTACAAGAC----------------------------------------- 2434

Qy    721 CCTACTCAGACTGTTACTTCTGGTGCACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 780
          -----------------------------------------------------------
Db   2434 ------------------------------------------------------------ 2434

Qy    781 AAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTTACCTACCTCATAGATTACTGCAACAG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2435 -----------------ACTCATAGATTACTGCAACAG 2455

Qy    841 TTCCCCCTGGACCTGAAAAGTTTCTTGCCTGCCTGCCTTACAGAGAGCTGAAACAACTGCCAAT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2456 TTCCCCCTGGACCTGAAAAGTTTCTTGCCTGCCTGCCTTACAGAGAGCTGAAACAACTGCCAAT 2515

Qy    901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGAGACTCCAAGGGAGTAAAAGAG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2516 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGAGACTCCAAGGGAGTAAAAGAG 2575

Qy    961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2576 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAC 2635

Qy   1021 AACCTGGATGAAAACAGCCAAAAAATCTGAGATCCCTGGAGGTTCCGATGATGCAGTC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2636 AACCTGGATGAAAACAGCCAAAAAAATCCTGGATCCCTGGAGGTTCCGATGATGCAGTC 2695

Qy   1081 CTGTTACAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2696 CTGTTACAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCT 2755

Qy   1141 CTCAACATTAGGTCCCATTTGGAGAGCACCAGTGGAAGCCGTCTGACCTTTCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2756 CTCAACATTAGGTCCCATTTGGAGAGCACCAGTGGAAGCCGTCTGACCCTTTCT 2815

Qy   1201 CTGCAGGAACTTCTGGTGTGCTACAGTGAAAGATGATGAATTAAGCCGGCAGGCACCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2816 CTGCAGGAACTTCTGGTGTGCTACAGTGAAAGATGATGAATTAAGCCGGCAGGCACCT 2875

Qy   1261 ATTGGAGGGCGACTTTCAGCAGTTCAGAAGCAGAAGCAGAACATGTACATGGAACAGGA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2876 ATTGGAGGGCGACTTTCAGCAGTTCAGAGAGCAGAAGCAGAACATGTACATGGAACAGGA 2935

Qy   1321 GAATTGAAAACTAAAGAACCTGTAATCATGAGGACTACTCTTGAGACTGTACGAATATTTCTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2936 GAATTGAAAACTAAAGAACCTGTAATCATGAGGACTACTCTTGAGACTGTACGAATATTTCTG 2995

Qy   1381 ACAGAGCAGCGCTTTGGAAGGACTAGAGAAACTTCACCAGGAGA 1440
          ||||||||||||||||||||||||||||||||||||||||||||
Db   2996 ACAGAGCAGCGCTTTGGAAGGACTAGAGAAACTTCACCAGGAGAGCTGCCTCCT 3055
```

```
Db  1840 GTATTATCGTGGCTTCTTTCTGCTGAGGAGCACATTGCAAGCACAAGGAGAGATTTCTAAT 1899
Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGATTGACA 180
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1900 GATGTGGAAGTGGTGAAAGACCAGTTTCATCTACTCATGAGGGGTACATGATGATTGACA 1959
Qy   181 GCCCATCAGGGCGCGGGTTGGTATATTCTACAATTGGGAGGTAAGCTGATTGGAACAGGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1960 GCCCATCAGGGCGCGGGTTGGTATATTCTACAATTGGGAGGTAAGCTGATTGGAACAGGA 2019
Qy   241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2020 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 2079
Qy   301 TGGGAATGCCTCAGGTAGCTAGCAGGTCATGGAAAAACAAAGCAATTACATAGAGTTTTAATG 360
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2080 TGGGAATGCCTCAGGTAGCTAGCAGGTCATGGAAAAACAAAGCAATTACATAGAGTTTTAATG 2139
Qy   361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2140 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 2199
Qy   421 ACAAGGAAAATGAGGAAGAGCCTCTTGGACTTGATCTTGAAGACCTAAAACGCCAAGTA 480
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2200 ACAAGGAAAATGAGGAAGAGCCTCTTGGACTTGATCTTGAAGACCTAAAACGCCAAGTA 2259
Qy   481 CAACACATAAGGTGCTTCAAGAAGATCTAGAACAAGCAAGTCAGGTCAATTCTCTC 540
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2260 CAACACATAAGGTGCTTCAAGAAGATCTAGAACAAGCAAGTCAGGTCAATTCTCTC 2319
Qy   541 ACTCACATGGTGTGGTAGTTGATGAATCTAGTGGACAACTCAGTGGATTGGAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2320 ACTCACATGGTGTGGTAGTTGATGAATCTAGTGGACAACTCAGTGGATTGGAA 2379
Qy   601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2380 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 2439
Qy   661 TGGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCTGACTCCTGGACTCATTGGAGCCTCT 720
         |||||||||||||||||||
Db  2440 TGGGGTTCTTTTACAAGAC 2457
Qy   721 CCTACTCAGACTGTTACTCTGGTGCACAACCTGTGGTTACTAAGGAAACTGCCACTCTCC 780
Db  2458 ------------ 2457
Qy   781 AAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
         ----------------------------ACTCATAGATTACTGCAACAG
Db  2458 ----ACTCATAGATTACTGCAACAG 2478
Qy   841 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 900
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2479 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAT 2538
Qy   901 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAG 960
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2539 GTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAG 2598
Qy   961 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGCTTTATCAC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2599 CTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACAGCTTTATCAC 2658
Qy  1021 AACCTGGATGAAACAGCCAAAAATCCTGAGATGCCTGGAAGGTTCCGATGATGCAGTC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2659 AACCTGGATGAAACAGCCAAAAATCCTGAGATGCCTGGAAGGTTCCGATGATGCAGTC 2718
Qy  1081 CTGTTACAAGACGTTGATAACATGAACAGCTTTAGAAGAA 1140
         ||||||||||||||||||||||||||||||||||||||||
Db  2719 CTGTTACAAGACGTTGATAACATGAACAGCTTTAGAAGAA 2778
Qy  1141 CTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTCGTCTGACCACCTTTTCT 1200
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2779 CTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTCGTCTGACCACCTTTTCT 2838
```

```
Qy  1441 GAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 1500
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3056 GAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 3115
Qy  1501 GAGTGGGAAAAATGAACCTGACTCCGCTGACTGGAGAGAAAAATAGATGAGACCCTT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3116 GAGTGGGAAAAATGAACCTGACTCCGCTGACTGGAGAGAAAAATAGATGAGACCCTT 3175
Qy  1561 GAAAGACTTCAGGAACTTCAAGAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 1620
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3176 GAAAGACTTCAGGAACTTCAAGAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCT 3235
Qy  1621 GAGGTGATCAAGGATCTCGGCGACCCGTGGGCAATCTCCATTGACTCTCTCCAAGAT 1680
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3236 GAGGTGATCAAGGATCTCGGCGACCCGTGGGCAATCTCCATTGACTCTCTCCAAGAT 3295
Qy  1681 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGC 1740
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3296 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGC 3355
Qy  1741 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTTCACCGTATAAC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3356 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTTCACCGTATAAC 3415
Qy  1801 CTCAGCACTTCTGAAGAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAGGAC 1860
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3416 CTCAGCACTTCTGAAGAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCGAGGAC 3475
Qy  1861 CGGAGTCAGGCAGCTGCATGAGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3476 CGGAGTCAGGCAGCTGCATGAGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTT 3535
Qy  1921 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTAT 1980
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3536 TCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTAT 3595
Qy  1981 ATCAACCACGAGACTCAAACA 2001
         |||||||||||||||||||||
Db  3596 ATCAACCACGAGACTCAAACA 3616
```

RESULT 6
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 85.4%; Score 1709; DB 10; Length 4848;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

```
Qy     1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1780 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAA 1839
Qy    61 GTATTATCGTGGCTTCTTTCTGCTGAGGACCACATTGCAAGCACAAGGAGAGATTTCTAAT 120
         ||||||||||||||||||||||||||||||||||||||||||||||
```

```
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match        85.4%;   Score 1709;   DB 10;   Length 5060;
Best Local Similarity   93.0%;   Pred. No. 0;
Matches 1860;   Conservative   0;   Mismatches   0;   Indels 141;   Gaps   1;

Qy      1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1992 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAA 2051

Qy     61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2052 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 2111

Qy    121 GATGTGGAAGTGGTGAAAGACCAGTTCATCATGTCATGATGGAGTATTGACA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2112 GATGTGGAAGTGGTGAAAGACCAGTTCATCATGTCATGATGGAGTATTGACA 2171

Qy    181 GCCCATCAGGGCCGGGTTGGTAATATTCACAATTGGGAGTAAGCTGATTGGAACAGGA 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2172 GCCCATCAGGGCCGGGTTGGTAATATTCACAATTGGGAGTAAGCTGATTGGAACAGGA 2231

Qy    241 AAATTATCAGAAGATGAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2232 AAATTATCAGAAGATGAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 2291

Qy    301 TGGGAATGCCTCAGGGTAGCTGGAAAAACTGAAAGAGTTGAATGACTGATCTGGCTAACAAAA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2292 TGGGAATGCCTCAGGGTAGCTGGAAAAACTGAAAGAGTTGAATGACTGATCTGGCTAACAAAA 2351

Qy    361 GATCTCCAGAATCAGAACTCAGAAACTGAAAGACCTCTTGAAGACCTAAAACGGCCAAGTA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2352 GATCTCCAGAATCAGAACTCAGAAACTGAAAGACCTCTTGAAGACCTAAAACGGCCAAGTA 2411

Qy    421 ACAAGAGAAAATGGAGAAGAGCCTCTTCAAGAAGATCTAGAACAAGAACAAGCAGGTCTCTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2412 ACAAGAGAAAATGGAGAAGAGCCTCTTCAAGAAGATCTAGAACAAGAACAAGCAGGTCTCTC 2471

Qy    481 CAACAACATAAGGTGCTTCAAGAGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCTCTC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2472 CAACAACATAAGGTGCTTCAAGAGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCTCTC 2531

Qy    541 ACTCACATGGTGGTGAGTTGATGAATCTAGTGGACAGAGACCGCTGCTTTGGAA 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2532 ACTCACATGGTGGTGAGTTGATGAATCTAGTGGACAGAGACCGCTGCTTTGGAA 2591

Qy    601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2592 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 2651

Qy    661 TGGGTTCTTTTACAAGACCAGCCTGACTGACCTCCTGGACTGACCTATTGGAGCCTCT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2652 TGGGTTCTTTTACAAGAC----------------------------------------- 2669

Qy    721 CCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCC 780
            ---------------------------------------------------------
Db   2670 ------------------------------------------------------ 2669

Qy    781 AAACTAGAAATGCCATCTTCCTTGATGTTGAGGTACCTACTCATAGATTACTGCAACAG 840
            ------------------------------------------ACTCATAGATTACTGCAACAG
Db   2670 ------------------------------------------ACTCATAGATTACTGCAACAG 2690

Qy    841 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACGAGAGCTGAAACAACTGCCAAT 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2691 TTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAAT 2750

Qy    901 GTCCTACAGGATGCTACCCGTAAGGAAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2751 GTCCTACAGGATGCTACCCGTAAGGAAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAG 2810
```

```
Qy   1201 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCAGCACCT 1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2839 CTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCAGCACCT 2898

Qy   1261 ATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGACGATGTACATAGGGCTTCAAGAGG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2899 ATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGACGATGTACATAGGGCTTCAAGAGG 2958

Qy   1321 GAATTGAAAACTAAAGAACCTGTAATCATGATGTCTTGAGACTGTACGAATATTTCTG 1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2959 GAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3018

Qy   1381 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTTCACCAGAGAGCCCAGAGCTGCCTCCT 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3019 ACAGAGCAGCCTTTGGAAGGACTAGAGAAACTTCACCAGAGAGCCCAGAGCTGCCTCCT 3078

Qy   1441 GAGGAGAGCAGCCCAGAGAATGTCACTCGGCTTCACTCGGCTTCTACGAAAGCAGGCTTCGAGAGGTCAATACT 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3079 GAGGAGAGCAGCCCAGAGAATGTCACTCGGCTTCACTCGGCTTCTACGAAAGCAGGCTTGAGAGGTCAATACT 3138

Qy   1501 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTCGGCGAGAGAAAAATAGATGAGACCCTT 1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3139 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTCGGCGAGAGAAAAATAGATGAGACCCTT 3198

Qy   1561 GAAAGACTTCCAGGAACTTCAAGAGGCCACGGACCTGAGCTGGACCTCAAGCTGCGCCAAGCT 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3199 GAAAGACTTCCAGGAACTTCAAGAGGCCACGGACCTGAGCTGGACCTCAAGCTGCGCCAAGCT 3258

Qy   1621 GAGGTGGATCAATGACCTTGCTCGGCCAGCACTTGAGAGAGAAAGTGAAGGCACTTCACCGTGAGC 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3259 GAGGTGGATCAATGACCTTGCTCGGCCAGCACTTGAGAGAGAAAGTGAAGGCACTTCACCGTGAGC 3318

Qy   1681 CACCTGCACTCTGAAGACTTGCTCGCCAGCTTACCACTTTGGGCCTTCGGAGAAGTTCTCACCGTATAAC 1740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3319 CACCTGCACTCTGAAGACTTGCTCGCCAGCTTACCACTTTGGGCCTTCGGAGAAGTTCTCACCGTATAAC 3378

Qy   1741 CACGTCACTCTGAAGAGCTTCGAGAGCTTCTGCAGGTGGCCGTCGAGGAC 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3379 CACGTCACTCTGAAGAGCTTCGAGAGCTTCTGCAGGTGGCCGTCGAGGAC 3438

Qy   1801 CTCAGCACTCTGAAGAGCTTCGAGAGCTTCTGCAGGTGGCCGTCGAGGAC 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3439 CTCAGCACTCTGAAGAGCTTCGAGAGCTTCTGCAGGTGGCCGTCGAGGAC 3498

Qy   1861 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGTCCAGCATCTCAGCACTTTCTT 1920
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3499 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGTCCAGCATCTCAGCACTTTCTT 3558

Qy   1921 TCCACGTCTGTCCAGGGGTCCCTGGGAGAGAGCCCACAAAAGTGCCCTACTAT 1980
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3559 TCCACGTCTGTCCAGGGGTCCCTGGGAGAGAGCCCACAAAAGTGCCCTACTAT 3618

Qy   1981 ATCAACCACGAGACTCAAACA 2001
            |||||||||||||||||||||
Db   3619 ATCAACCACGAGACTCAAACA 3639
```

RESULT 7
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1

RESULT 8
US-09-845-416-2
; Sequence 2, Application US/0984516
; Publication No. US2003017131A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match   83.1%;   Score 1662.8;   DB 10;   Length 4182;
Best Local Similarity   88.3%;   Pred. No. 0;
Matches 1929;   Conservative   0;   Mismatches   72;   Indels   183;   Gaps   4;

```
Qy       1 GGCAGTTCATTGATGATGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA    60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1000 GGCAGTTCATTGATGATGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA  1059

Qy      61 GTATTATCGTGGCTTCTTTCGTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1060 GTATTATCGTGGCTTCTTTCGTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT  1119

Qy     121 GATGTGGAAGTGTGAAAGACCAGTTCATACTCATGAGGGGTACATGAGTGATTGACA   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1120 GATGTGGAAGTGTGAAAGACCAGTTCATACTCATGAGGGGTACATGAGTGATTGACA  1179

Qy     181 GCCCATCAGGGCGCGGGTTGGTATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1180 GCCCATCAGGGCGCGGGTTGGTATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA  1239

Qy     241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGA   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1240 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGCAGATGAATCTCCTAAATTCAAGA  1299

Qy     301 TGGGAATGCCTCAGGGTAGCTAGCAGCATGCATGGAAAAACAAAGCAATTTTACATAGAGTTTTAATG   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1300 TGGGAATGCCTCAGGGTAGCTAGCAGCATGCATGGAAAAACAAAGCAATTTTACATAGAGTTTTAATG  1359

Qy     361 GATCTCCAGAATCAGAAACTAGAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1360 GATCTCCAGAATCAGAAACTAGAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA  1419

Qy     421 ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACTAAAACGCCAAGTA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1420 ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACTAAAACGCCAAGTA  1479

Qy     481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAAGTCAGGTCAATTCTCTC   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1480 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAAGTCAGGTCAATTCTCTC  1539

Qy     541 ACTCACATGGTGTGTGGTAGTTGATGAATCTAGTAGGAACTAGTGTGCTTGCTTTGGAA   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1540 ACTCACATGGTGTGTGGTAGTTGATGAATCTAGTAGGAACTAGTGTGCTTGCTTTGGAA  1599

Qy     601 GAACAACTTAAGGTATTGGGAGATCAATGGGCAAAACATCTGTAGTGGACGAAGACCGC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1600 GAACAACTTAAGGTATTGGGAGATCAATGGGCAAAACATCTGTAGTGGACGAAGACCGC  1659

Qy     661 TGGGTTCTTTACAAGACCAGCCGAC-------------CTAGTCCTGGACTGACCACT   708
           ||||||||||||||||||||||||||             ||||||||||||||||||
Db    1660 TGGGTTCTTTTACAAGACATCCTTCAAATGGCAACGTCTCAAATGGCAACGTCTTACTGGAACGTGAAGAACAGAGTGCCTT  1719
```

Qy   709 ATTGGAGCCTCCTACTCAGACTGTTACTCTGTGACACA--------ACCTGTCGGTT 759
          ||||||| ||||| ||||| | ||| | | | ||||| ||||
Db 1720 TTTAGTGCATGGCTTTCAGAAAAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTT 1779

Qy   760 ACTAAGGAAACTGCCATCTC----------------------------------- 779
          || |||| |||| ||||||
Db 1780 AAAGATCAAAATGAAAATGTTATCAAGTCTTCAAAAACTCGCCGTTTAAAAGCGGATCTA 1839

Qy   780 --------------------------------CAAACTAGAAATGCCATCTTCC 801
                                          ||||| |||||| ||||||||
Db 1840 GAAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACA 1899

Qy   802 TTGATGTTGGAG------------ 813
          ||| |||||
Db 1900 CTGAAGAATAAGTCAGTGACCAGAGACGGAAGCATGGCTGGATAACTTGCCCGGTGT 1959

Qy   814 -------------------------GTACCTACTCATAGATTACTGCAA 837
                                   |||||||||||||||||||||||||
Db 1960 TGGGATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2019

Qy   838 CAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGTGAAACAACTGCC 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2020 CAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGTGAAACAACTGCC 2079

Qy   898 AATGTCCTACAGGATGCTACCCGTAAGGAAAAGGCTTCCTAGAAGACTTCCAAGGGAGTAAAA 957
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db 2080 AATGTCCTACAGGATGCTACCCGTAAGGAAAAGGCTTCCTAGAAGACTTCCAAGGGAGTAAAA 2139

Qy   958 GAGCTGATGAAACAATGGCAAGACCTTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 1017
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2140 GAGCTGATGAAACAATGGCAAGACCTTCCAAGGTGAAATTGAAGCTCACACAGATGTTTAT 2199

Qy 1018 CACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 1077
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2200 CACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 2259

Qy 1078 GTCCTGTACAAAGACGTTTGGATAACATGACTTCAAGTGGAGTGAACTTCGGAAAAAG 1137
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2260 GTCCTGTACAAAGACGTTTGGATAACATGACTTCAAGTGGAGTGAACTTCGGAAAAAG 2319

Qy 1138 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCGACCAGTGGAAGGGTCTGCACCTT 1197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2320 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCGACCAGTGGAAGGGTCTGCACCTT 2379

Qy 1198 TCTCTGCAGGAACTTCTGGTGTGTGGCTACAGCTGAAAGATGAAGATTAAGCCCGGCAGGCA 1257
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2380 TCTCTGCAGGAACTTCTGGTGTGTGGCTACAGCTGAAAGATGAAGATTAAGCCCGGCAGGCA 2439

Qy 1258 CCTAATTGGAGGCGACTTCCAGCAGTTCAGAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTT 1317
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2440 CCTAATTGGAGGCGACTTCCAGCAGTTCAGAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTT 2499

Qy 1318 AGGGAATTGAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTT 1377
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2500 AGGGAATTGAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTT 2559

Qy 1378 CTGAGACAGAGCAGCCCAGAGAGACTAGAGAAACTCTACCAGGAGCCCAGAGCTGCCT 1437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2560 CTGACAGAGCAGCCCAGAGAGACTACCAGGAGAAACTCTACCAGGAGCCCAGAGCTGCCT 2619

Qy 1438 CCTGAGGAGAGACCCAGAGATGTCACTCGGCTTCACGGAAAGCAGGCTGAGGAGGTCAAT 1497
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2620 CCTGAGGAGAGACCCAGAGATGTCACTCGGCTTCACGGAAAGCAGGCTGAGGAGGTCAAT 2679

Qy 1498 ACTGAGTCGGAAAAATTGAACCTGACTGACTCGGCTGGCAGAGAAAAATAGATGAAGAGACC 1557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2680 ACTGAGTCGGAAAAATTGAACCTGACTGACTCGGCTGGCAGAGAAAAATAGATGAAGAGACC 2739

Qy 1558 CTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGGTGCGCGCCAA 1617
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2740 CTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGGTGCGCGCCAA 2799

Qy 1618 GCTGAGGTGATCAAGGATCCTGGCCAGCCCGTGGGCGATCCTCCTCATTGACTCTCTCCAA 1677
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2800 GCTGAGGTGATCAAGGATCCTGGCCAGCCCGTGGGCGATCCTCCTCATTGACTCTCTCCAA 2859

Qy 1678 GATCACCTCGAGAAGTCAAGGCACTTCGAGGAGAAATTGCCCTGAAAGAGAACGTG 1737
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2860 GATCACCTCGAGAAGTCAAGGCACTTCGAGGAGAAATTGCCCTGAAAGAGAACGTG 2919

Qy 1738 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTAT 1797
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2920 AGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTAT 2979

Qy 1798 AACCTCAGCACTCTGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAG 1857
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2980 AACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCCGTCGAG 3039

Qy 1858 GACGAGTCAGGCAGCTGCATGAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTT 1917
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3040 GACGAGTCAGGCAGCTGCATGAGCCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTT 3099

Qy 1918 CTTTCCACGTCGTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTAC 1977
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3100 CTTTCCACGTCGTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTAC 3159

Qy 1978 TATATCAACCACGAGACTCAAACA 2001
          ||||||||||||||||||||||||
Db 3160 TATATCAACCACGAGACTCAAACA 3183

RESULT 9
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27

Query Match   83.1%; Score 1662.8; DB 10; Length 5149;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

Qy   1 GGCAGTTCATTGATGGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1757 GGCAGTTCATTGATGGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 1816

Qy   61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 1876

Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGATGATGATTGACA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1877 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGATGATGATTGACA 1936

Qy   181 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGAAGTGAATTGGAAGCTGATTGAACAGGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1937 GCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGAAGTGAATTGGAAGCTGATTGAACAGGA 1996

Qy   241 AAATTATCAGAAGATGAAGAAACTGAAGTAAGTAGATACAAGAGCAGAGAATCTCCTAAATTCAAGA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1997 AAATTATCAGAAGATGAAGAAACTGAAGTAAGTAGATACAAGAGCAGAGAATCTCTAAATTCAAGA 2056

Qy   301 TGGGAATGCCTCCAGGGTAGCTAGCATGGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2057 TGGGAATGCCTCCAGGGTAGCTAGCATGGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATG 2116

Qy   361 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2117 GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGA 2176

Qy   421 ACAAGGAAAATGAGGAGGAGAGCGCTCTTGAAGACCTGAAGACCTAAAACGCCAAGTA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2177 ACAAGGAAAATGAGGAGGAGAGCGCTCTTGAAGACCTGAAGACCTAAAACGCCAAGTA 2236

Qy   481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2237 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 2296

Qy   541 ACTCACGATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2297 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2356

Qy   601 GAACAACTTAAGGTATTGGGAATCGATGGGCCAAACATCTGTAGATGGACAGAAGACCGC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2357 GAACAACTTAAGGTATTGGGAATCGATGGGCCAAACATCTGTAGATGGACAGAAGACCGC 2416

Qy   661 TGGGTTCTTTTACAAGACCAGCCTGAC-------CTAGCTCCTGGACTGACCACT 708
           ||||||||||||||||||||||||||| |||||||||||||||||||||
Db   2417 TGGGTTCTTTTACAAGACCAGCCTGACCATCCTTCTCAAATGGCAACGTCTTTACTGGAACAGTGCCTT 2476

Qy   709 ATTGGAGCCTCTCCACTCAGACTGTTACTCTGGTGACACA-------ACCTGTGGTT 759
           |||||||||||||||||||||||||||||||||| | |||||||
Db   2477 TTTAGTGCATGGCTTTCAGAAAAAAGAAGAGCAGTGAACAGAGATTCACACAACAACTGGCTTT 2536

Qy   760 ACTAAGGAAACTGCCATCTC---------- 779
           ||||||||||||||||||
Db   2537 AAAGATCAAAATGAAAATGTTATCAAGTCTTATCAAGTGCTTTAAAAGGCCGATCTA 2596

Qy   780 -----------CAAAACTAGAAATGCCATCTTCC 801
                     ||||| ||||||||||||||||
Db   2597 GAAAAGAAAAAGCAAATCCATGGGCAAACTGTATTCACTCAAACAGATCTTCTTCTTCAACA 2656

Qy   802 TTGAGTTGGAG------------- 813
           ||||||||||
Db   2657 CTGAAGAATAAGTCAGTGACCCAGAGACGGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGT 2716

Qy   814 ------------------------GTACCTACTCATAGATTACTGCAA 837
                                   ||||||||||||||||||||||||
Db   2717 TGGGATAATTTAGTCCAAAAAACTTGAAAAGACTACAGACAGCTCATAGATTACTGCAA 2776

Qy   838 CAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCC 897
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2777 CAGTTCCCCTGTGTCCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAAACAACTGCC 2836

Qy   898 AATGTCCTACAGGAATGCTACCCGTAGGAAAAGGCTCCTAGAAGAGACTCTCCAAGGGAGTAAAA 957
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2837 AATGTCCTACAGGATGCTACCCGTAGGAAAAGGCTCCTAGAAGAGACTCTCCAAGGGAGTAAAA 2896

Qy   958 GAGCTGATGAAACAATGCAAGTCCAAGAGTGAAATTGAAGTCCACACAGATGTTTAT 1017
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2897 GAGCTGATGAAACAATGGCCAAGAGTCCAAGAGTGAAATTGAAGTCCACACAGATGTTTAT 2956

Qy   1018 CACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 1077
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2957 CACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 3016

Qy   1078 GTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAG 1137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3017 GTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAG 3076

Qy   1138 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTGGTGAAGCCTGTCGACCTT 1197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3077 TCTCTCAACATTAGGTCCCATTTGGAAGCCAGTGGTGAAGCCTGTCGACCTT 3136

Qy   1198 TCTCTGCAGGAACTTCTGGTGTGGCCTACAGCTGAAAGATGATGATGATTTTTT|||||||||||||AAGCCGGCAGGCA 1257
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3137 TCTCTGCAGGAACTTCTGGTGTGGCCTACAGCTGAAAGATGATGATGAATTAAGCCGGCAGGCA 3196

Qy   1258 CCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACAGATGTACATAGGGCCTTCAAG 1317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3197 CCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACAGATGTACATAGGGCCTTCAAG 3256

Qy   1318 AGGGAATGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAACGAATATTT 1377
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3257 AGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAACGAATATTT 3316

Qy   1378 CTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGAAACTACCAGGAGCCCAGACTGCCT 1437
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3317 CTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGAAACTACCAGGAGAGCTGCCT 3376

Qy   1438 CCTGAGAGAGCCCAGAATGTCACTCGGCTTCACGAAAGTCGAGGAGGTCAAT 1497
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3377 CCTGAGGAGAGCCCAGAAATGTCACTCGGCTTCACGAAAGTCGAGGAGGTCAAT 3436

Qy   1498 ACTGAGTGGGAAAAATTGAAACTGCACTCGCACTCGCACTGGCCAGAGAAAAATGATGAGACC 1557
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3437 ACTGAGTGGGAAAAATTGAAACTGCACTCGCACTCCCGCTGGCCAGAGAAAAATAGATAGACC 3496

Qy   1558 CTTGAAAGACTCCAGGAACTTCAGGAGCTGGCACCTCAAGCTGCGCCAA 1617
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3497 CTTGAAAGACTCCAGGAACTTCAGGAGCTGGCACCTCAAGCTGCGCCAA 3556

Qy   1618 GCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGGCGATCTTCCTCATTGACTCTCTCCAA 1677
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3557 GCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAA 3616

Qy   1678 GATCACCTTCGAGAAGTCAAGGCACTTCGAGAGAAATTGCGCCTCTGAAAGAGACGTG 1737
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3617 GATCACCTTCGAGAAGTCAAGGCACTTCGAGAGAAATTGCGCCTCTGAAAGAGAACGTG 3676

Qy   1738 AGCCACGTCAATGACCTTGCTCGCGAGCTTACCACTTTGGGCATTCAGCTCTCACCGTAT 1797
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3677 AGCCACGTCAATGACCTTGCTCGCGAGCTTACCACTTTGGGCATTCAGCTCTCACCGTAT 3736

Qy   1798 AACCTCAGCACTCTGGAAGACCTGACACCAGATGGAAGCTTCTGCGGCTTGGCCGTCGAG 1857
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3737 AACCTCAGCACTCTGGAAGACCTGACACCAGATGGAAGCTTCTGCGGCTTGGCCGTCGAG 3796

Qy   1858 GACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGTCGTCCAGCATCTCAGCACTTT 1917
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3797 GACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGTCGTCCAGCATCTCAGCACTTT 3856

Qy   1918 CTTTCACGTCTGTCCAGGGTCCCTGGCAAATGGCCATCTCGCCAAACAAAGTGCCCTAC 1977
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3857 CTTTCCACGTCTGTCCAGGGTCCTGGCAAATGGCCATCTCGCCAAACAAAGTGCCCTAC 3916

Qy   1978 TATATCAACCACGAGACTCAAACA 2001
           ||||||||||||||||||||||||
Db   3917 TATATCAACCACGAGACTCAAACA 3940

RESULT 10
US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US2003021633A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRNT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/235,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 5462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-41

Query Match      64.1%;   Score 1283;   DB 15;   Length 5462;
Best Local Similarity   99.3%;   Pred. No. 0;
Matches 1302;   Conservative   0;   Mismatches   0;   Indels   9;   Gaps   1;

Qy   691 GCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAA 750
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1547 GCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAA 1606

Qy   751 CCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTG 810
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1607 CCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTG 1666

Qy   811 GAGGTACCTACTCATAGATTACTGCAACAGTCCCCCTGGACCTGGAAAAGTTTCTTGCC 870
         |||       ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1667 GAG-------CATAGATTACTGCAACAGTCCCCCTGGACCTGGAAAAGTTTCTTGCC 1717

Qy   871 TGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACACGATGCTACCCGTAAGGAAAGG 930
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1718 TGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACACGATGCTACCCGTAAGGAAAGG 1777

Qy   931 CTCCTAAGAGACTCCAAGGAGTAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGT 990
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1778 CTCCTAAGAGACTCCAAGGAGTAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGT 1837

Qy   991 GAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTG 1050
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1838 GAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTG 1897

Qy  1051 AGATCCCTGGAAGGTTCCGATGATGCAGTCCTGGTTACAAAGACGTTTGGATAACATGAAC 1110
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1898 AGATCCCTGGAAGGTTCCGATGATGCAGTCCTGGTTACAAAGACGTTTGGATAACATGAAC 1957

Qy  1111 TTCAAGTGGAGTGAACTTCGGGAAAAAGTCTCTCAACATTGGTTCCCATTTGGAAGCCAGT 1170
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1958 TTCAAGTGGAGTGAACTTCGGGAAAAAGTCTCTCAACATTGGTTCCCATTTGGAAGCCAGT 2017

Qy  1171 TCTGACCAGTGGAAGCGTCTGCACCTTTGCTGGTGTGGCTCACAGCTG 1230
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2018 TCTGACCAGTGGAAGCGTCTGCACCTTTGCTGGTGTGGCTCACAGCTG 2077

Qy  1231 AAAGATGATGAATTAAGCCGGCAGCCACCTATTGGAGGGGACTTTCAGCAGTTCAGAAG 1290
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2078 AAAGATGATGAATTAAGCCGGCAGCCACCTATTGGAGGGGACTTTCAGCAGTTCAGAAG 2137

Qy  1291 CAGAACGATGTACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAACCTGTAATCATG 1350
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2138 CAGAACGATGTACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAACCTGTAATCATG 2197

Qy  1351 AGTACCTACTCATAGATATTCTGACAGAGACTGCCTTGGAAGGACTAGAGAAA 1410
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2198 AGTACCTACTCATAGATATTCTGACAGAGACTGCCTTGGAAGGACTAGAGAAA 2257

Qy  1411 CTCTACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTT 1470
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2258 CTCTACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTT 2317

Qy  1471 CTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCT 1530
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2318 CTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCT 2377

Qy  1531 GACTGCAGAGAAAAATAGATGAGACTCCAGGAACTTCAGAGGCCACG 1590
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2378 GACTGCAGAGAAAAATAGATGAGACTCCAGGAACTTCAGAGGCCACG 2437

Qy  1591 GATGAGCTGAACCTCAAGGTCGCGCCAAGCTGAGCGTGATCAAGGATCCTGGCAGCCCGTG 1650
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2438 GATGAGCTGAACCTCAAGGTCGCGCCAAGCTGAGCGTGATCAAGGATCCTGGCAGCCCGTG 2497

Qy  1651 GGCGATCTCCTCATTGACTCTCCAAGATCACCTGCAAGAAAGTCAAGGACCACTTCGAGGA 1710
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2498 GGCGGATCTCCTCATTGACTCTCCAAGATCACCTGCAAGAAAGTCAAGGACCACTTCGAGGA 2557

Qy  1711 GAAATTGCGCCTGAAAGAGACGTGAGCCACGCAGTGAGCCACGCAGTGAGCCAGCTACC 1770
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2558 GAAATTGCCGCCTGAAAGAGACGTGAGCCACGCAGTGAGCCACGCAGTGAGCCAGCTACC 2617

Qy  1771 ACTTTGGCATTCAGCTCTCACCGTATAACCTCGGAAGACCTGAAGACACTCGAACACCAGA 1830
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2618 ACTTTGGGCATTCAGCTCTCACCGTATAACCTCGGAAGACCTGAAGACACTCGAACACCAGA 2677

Qy  1831 TGGAAGCTTCTGCAGGTGGCCGTCGAGGACGAGTCAGGCAGCTGCATGAGAGCCCACAGG 1890
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2678 TGGAAGCTTCTGCAGGTGGCCGTCGAGGACGAGTCAGGCAGCTGCATGAGAGCCCACAGG 2737

Qy  1891 GACTTTGGTCAGCATCTCAGCACTTCTTCCACGTCGTGTCCAGGGTCCCTGGGAGAGA 1950
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2738 GACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCGTGTCCAGGGTCCCTGGGAGAGA 2797

Qy  1951 GCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2798 GCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2848
```

RESULT 11

```
US-10-149-736-42
; Sequence 42, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-42

Query Match      59.1%;   Score 1182.6;   DB 15;   Length 8689;
Best Local Similarity   99.7%;   Pred. No. 0;   Gaps   0;
Matches 1185;   Conservative   0;   Mismatches   4;   Indels   0;

Qy   813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAGTTTCTTGCCTG 872
         ||                        ||||||||||||||||||||||||||||||||
Db  2992 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAGTTTCTTGCCTG 3051

Qy   873 GCTTACAGAGTCGAAACAACTGCCAATGTCCTACAGGATGCTCCCGTAAGGAAAGGCT 932
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3052 GCTTACAGAGTCGAAACAACTGCCAATGTCCTACAGGATGCTCCCGTAAGGAAAGGCT 3111

Qy   933 CCTAGAAGACTTCCAAGGAGTAAAAGAGTGATGAAACAATGGCAAGACCTCCAAGGTGA 992
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3112 CCTAGAAGACTTCCAAGGAGTAAAAGAGTGATGAAACAATGGCAAGACCTCCAAGGTGA 3171

Qy   993 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAG 1052
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3172 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAG 3231
```

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-1

Query Match   59.1%; Score 1182.6; DB 10; Length 11058;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy     813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 872
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8052 GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG 8111

Qy     873 GCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGTACCCGTAAGGAAAGGCT 932
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8112 GCTTACAGAGCTGAAACAACTGCCAATGTCCTACAGGATGTACCCGTAAGGAAAGGCT 8171

Qy     933 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGGTGA 992
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8172 CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA 8231

Qy     993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAAACAGCCAAAAAATCCTGAG 1052
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8232 AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG 8291

Qy    1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGATAACATGAACTT 1112
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8292 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGATAACATGAACTT 8351

Qy    1113 CAAGTGGAGTGAACTTGGAAAAGTCTTCTCAACATTGGAAGCCAGTTC 1172
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    8352 CAAGTGGAGTGAACTTGGAAAAGTCTTCTCAACATTGGAAGCCAGTTC 8411

Qy    1173 TGACCAGTGGAAGGTCTGCACCTTCTGCGAGGAACTTCGTGGTGCTACAGCTGAA 1232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8412 TGACCAGTGGAAGGTCTGCACCTTCTGCGAGGAACTTCGTGGTGCTACAGCTGAA 8471

Qy    1233 AGATGATGAAATAAGCCGGCACCCATTGGAGGCCAGTTCAGAAGCA 1292
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    8472 AGATGATGAAATAAGCCGGCACCCATTGGAGGCCAGTTCAGAAGCA 8531

Qy    1293 GAACGATGTACATAGGGCCTTCAAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8532 GAACGATGTACATAGGGCCTTCAAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 8591

Qy    1353 TACTCTTGAGACTGTACGAATATTTCTGACAAGCAGCCTTTGGAAGGACTAGAGAAACT 1412
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8592 TACTCTTGAGACTGTACGAATATTTCTGACAAGCAGCCTTTGGAAGGACTAGAGAAACT 8651

Qy    1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGAGAGGAATGTCACTCGGCTTCT 1472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8652 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGAGAGGAATGTCACTCGGCTTCT 8711

Qy    1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 1532
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8712 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA 8771

Qy    1533 CTGGCAGGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGGCCACGGA 1592
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8772 CTGGCAGGAAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGGCCACGGA 8831

Qy    1593 TGAGCTGGACCTCAAGCTGCGCCAAGTGAGTGAGCTGAATCAAGGATCCTGGG 1652
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8832 TGAGCTGGACCTCAAGCTGCGCCAAGTGAGTGAGCTGAATCAAGGATCCTGGG 8891
```

```
Qy    1053 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGATAACATGAACTT 1112
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3232 ATCCCTGGAAGGTTCCGATGATGCAGTCTCTGTTACAAAGACGTTTGATAACATGAACTT 3291

Qy    1113 CAAGTGGAGTGAACTTCGGAAAAAGTCTCCATTTGGTGTGGCTACAGCTGAA 1172
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3292 CAAGTGGAGTGAACTTCGGAAAAAGTCTCCATTTGGTGTGGCTACAGCTGAA 3351

Qy    1173 TGACCAGTGGAAGCCGTCGCACCTTTCTCGCGAGGAACTTCTGGTGTGGCTACAGCTGAA 1232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3352 TGACCAGTGGAAGCCGTCGCACCTTTCTCGCGAGGAACTTCTGGTGTGGCTACAGCTGAA 3411

Qy    1233 AGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAGCA 1292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3412 AGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAGCA 3471

Qy    1293 GAACGATGTACATAGGGGCCTTCAAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3472 GAACGATGTACATAGGGGCCTTCAAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 3531

Qy    1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCAGGGCCCAGAGAAACT 1412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3532 TACTCTTGAGACTGTACGAATATTTCTGACGACGAGCAGCCAGGGCCCAGAGAAACT 3591

Qy    1413 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCT 1472
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3592 CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCT 3651

Qy    1473 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATGAACCTGCACTCCGCTGA 1532
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3652 ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATGAACCTGCACTCCGCTGA 3711

Qy    1533 CTGGCAGAGAAAAATAGAGACTCCAGGAACTTCAAGAGAGGGCCACCGGA 1592
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    3712 CTGGCAGAGAAAAATAGAGACTCCAGGAACTTCAAGAGAGGGCCACCGGA 3771

Qy    1593 TGAGCTGGACCTCAAGCTGCGCCAAGTGAGGTGATCAAGGATCTGGCAGCCCGTGGG 1652
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3772 TGAGCTGGACCTCAAGCTGCGCCAAGTGAGGTGATCAAGGATCTGGCAGCCCGTGGG 3831

Qy    1653 CGATCTCCTCATTGACTTCTCCAAGATCACCTGAGAAAGTCAAGGCACTTCGAGGAGA 1712
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3832 CGATCTCCTCATTGACTTCTCCAAGATCACCTGAGAAAGTCAAGGCACTTCGAGGAGA 3891

Qy    1713 AATTGCGCCTCTGAAAAGAAGAACGTGAGCCACGCACGTCAATGACCTTGCTCGCCAGCTTACCAC 1772
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3892 AATTGCGCCTCTGAAAAGAAGAACGTGAGCCACGCACGTCAATGACCTTGCTCGCCAGCTTACCAC 3951

Qy    1773 TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCCACTCTGGAAGACCTGAACACCAGATG 1832
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3952 TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCCACTCTGGAAGACCTGAACACCAGATG 4011

Qy    1833 GAAGCTTCTGCAGGTGGGCCGTGGGCGTGGGGCACGACCGAGTCAGGCGACCGAGTCGCATGAAGCCCAGGGA 1892
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4012 GAAGCTTCTGCAGGTGGGCCGTGGGCGTGGGGCACGACCGAGTCAGGCGACCGAGTCGCATGAAGCCCAGGGA 4071

Qy    1893 CTTTGGTCCAGCATCTCAGCACTTCTTTCTTCCCACGTCGTGTCCAGGGGTCCTGGGAGAGAGC 1952
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4072 CTTTGGTCCAGCATCTCAGCATCTCAGCACTTCTTTCTTCCCACGTCGTGTCCAGGGGTCCTGGGAGAGAGC 4131

Qy    1953 CATCTCGCCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 2001
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4132 CATCTCGCCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 4180
```

RESULT 12
US-09-845-416-1
; Sequence 1, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

```
Db   5986  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT  6045

Qy   1113  CAAGTTGGAGTGAACTTCGAAAAAAAGTCTTCAACATTAGTCCATTGGAGCCAGTTC     1172
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6046  CAAGTTGGAGTGAACTTCGAAAAAAAGTCTTCAACATTAGTCCATTGGAGCCAGTTC     6105

Qy   1173  TGACCAGTGGAAGGCGTCTGCCACCTTTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA  1232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6106  TGACCAGTGGAAGGCGTCTGCCACCTTTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAA  6165

Qy   1233  AGATGATGAATTAAGCCGGCAGGCACCATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA   1292
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6166  AGATGATGAATTAAGCCGGCAGGCACCATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA   6225

Qy   1293  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6226  GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  6285

Qy   1353  TACTCTTGAGACTGTACGAATATTTCTGACAGAGGACTAGAGAAACT               1412
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   6286  TACTCTTGAGACTGTACGAATATTTCTGACAGAGGACTAGAGAAACT               6345

Qy   1413  CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAAATTGAACCTGCACTCCGCTGA  1472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6346  CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAAATTGAACCTGCACTCCGCTGA  6405

Qy   1473  ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA  1532
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6406  ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA  6465

Qy   1533  CTGGCAGAGAAAATAGATGAAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGCCACGGA  1592
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6466  CTGGCAGAGAAAATAGATGAAGACCCTTGAAAGACTTCAGGAACTTCAAGAGGCCACGGA  6525

Qy   1593  TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGG   1652
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6526  TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGG   6585

Qy   1653  CGATCTCCTCATTGACTCTCTCCAAGATCACCTGCAGAAAGTCAAGGCACTTCGAGGAGA  1712
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6586  CGATCTCCTCATTGACTCTCTCCAAGATCACCTGCAGAAAGTCAAGGCACTTCGAGGAGA  6645

Qy   1713  AATTGGGCCTCTGAAAGAGAACGTGAGCGAGCACGTTCACCGTATAAACCTGAACACCAGATG  1772
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6646  AATTGGGCCTCTGAAAGAGAACGTGAGCGAGCACGTTCACCGTATAAACCTGAACACCAGATG  6705

Qy   1773  TTTGGGCATTCAGCTCTCACCGTTCCCAGGTCCCTGTCCAGCGTCTGTCCCAGGGTCCCTG  1832
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6706  TTTGGGCATTCAGCTCTCACCGTTCCCAGGTCCCTGTCCAGCGTCTGTCCCAGGGTCCCTG  6765

Qy   1833  GAAGCTTCTGCAGGTGCCGTGAGGACCAGGTGCAGCTGCGATGAGAAGCCCACAGGGA   1892
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6766  GAAGCTTCTGCAGGTGCCGTGAGGACCAGGTGCAGCTGCGATGAGAAGCCCACAGGGA   6825

Qy   1893  CTTTGGTCCAGCATCCAGCACTTCTTCGACACTCAGGTCCCTGTCCAGGGTCCCTGGGAGAGC  1952
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6826  CTTTGGTCCAGCATCCAGCACTTCTTCGACACTCAGGTCCCTGTCCAGGGTCCCTGGGAGAGC  6885

Qy   1953  CATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACA              2001
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   6886  CATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA             6934
```

RESULT 13
US-10-149-736-44
; Sequence 44, Application US/10149736
; Publication No. US2003021633A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 11443
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-44

Query Match   59.1%; Score 1182.6; DB 15; Length 11443;
Best Local Similarity   99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy    813  GGTACCTACTCATAGATTACTGCAAGAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG   872
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5746  GGAAGAAACTCATAGATTACTGCAAGAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTG   5805

Qy    873  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT   932
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5806  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT   5865

Qy    933  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA   992
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5866  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA   5925

Qy    993  AATTGAAGCTCACACAGATGTTTATCACAACAACTGGATGAAAAACAGCCAAAAATCCTGAG  1052
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5926  AATTGAAGCTCACACAGATGTTTATCACAACAACTGGATGAAAAACAGCCAAAAATCCTGAG  5985

Qy   1053  ATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTT  1112
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 14
US-10-149-736-47
; Sequence 47, Application US/10149736
; Publication No. US2003021633A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736

; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 12057
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-47

Query Match   59.1%;   Score 1182.6;   DB 15; Length 12057;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy   813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTG   872
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8260  GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTG  8319

Qy   873  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGTCTACAGGATGCTAAGGCT   932
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8320  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGTCTACAGGATGCTAAGGCT  8379

Qy   933  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA   992
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8380  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA  8439

Qy   993  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG  1052
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8440  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG  8499

Qy  1053  ATCCCTGGAAGTTCCGATGATGCAGTCCGTTTACAAAGACGTTTGGATAACATGAACTT  1112
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8500  ATCCCTGGAAGTTCCGATGATGCAGTCCGTTTACAAAGACGTTTGGATAACATGAACTT  8559

Qy  1113  CAAGTGGAGTGAACTTCGAGGAACTTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  1172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8560  CAAGTGGAGTGAACTTCGAGGAACTTCTCAACATTAGGTCCCATTTGGAAGCCAGTTC  8619

Qy  1173  TGACCAGTGGAAGGGTCTGCACCTTCGTGGTGGCTACAGCTGAA  1232
          ||||||||||||||||||||||||||||||||||||||||||||
Db  8620  TGACCAGTGGAAGGGTCTGCACCTTCGTGGTGGCTACAGCTGAA  8679

Qy  1233  AGATGATGAATTAAGCCGGCAGCCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA  1292
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8680  AGATGATGAATTAAGCCGGCAGCCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA  8739

Qy  1293  GAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8740  GAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG  8799

Qy  1353  TACTCTTGAGACTGTACGAATATTTCTGACAGAGGACTAGAGAAACT  1412
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  8800  TACTCTTGAGACTGTACGAATATTTCTGACAGAGGACTAGAGAAACT  8859

Qy  1413  CTACCAGGAGCAGGCCCAGAGAGAGAGCCAGAATGTCACTCGGCTTCT  1472
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  8860  CTACCAGGAGCAGGCCCAGAGAGAGAGCCAGAATGTCACTCGGCTTCT  8919

Qy  1473  ACGAAAGCAGGCTGGCTGAGGTCAATACTGAGTGGGAAAAATTGAACCTGA  1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8920  ACGAAAGCAGGCTGGCTGAGGTCAATACTGAGTGGGAAAAATTGAACCTGA  8979

Qy  1533  CTGCAGAGAAAAATAGATGAGACCCCTGAAAGACTTCAAGAGGGCCACGGA  1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8980  CTGGCAGAGAAAAATAGATGAGACCCCTGAAAGACTTCAAGAGGGCCACCGGA  9039

Qy  1593  TGAGCTGGACCTCAAGCTGCGCGCCAAGCTGCGATGGATCAAGGGATCGCCGCAGCCCGTGGG  1652
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9040  TGAGCTGGACCTCAAGCTGCGCGCCAAGCTGCGATGGATCAAGGGATCGCCGCAGCCCGTGGG  9099

Qy  1653  CGATCTCCTCCATTGACTGCTCTCCAAGATCACCTGGAGAAAGTCAAGGCACTTCGAGGAGA  1712
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9100  CGATCTCCTCCATTGACTGCTCTCCAAGATCACCTGGAGAAAGTCAAGGCACTTCGAGGAGA  9159

Qy  1713  AATTGCGGCCTCTGAAAGAGAACGTGAGCGACGTGAGCGACGTGCTCCGCCAGCTTACCAC  1772
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9160  AATTGCGGCCTCTGAAAGAGAACGTGAGCGACGTGAGCGACGTGCTCCGCCAGCTTACCAC  9219

Qy  1773  TTTGGGCATTCAGCTCCACGTATAACCTGCACTCTGGAAGACTCTGGAAGACACCAGATG  1832
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9220  TTTGGGCATTCAGCTCCACGTATAACCTGCACTCTGGAAGACTCTGGAAGACACCAGATG  9279

Qy  1833  GAAGCTTCTGCAGTGGCCGTGAGGACCAGCAGTGAGGACCAGCATGAAGCCCACAGGGA  1892
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9280  GAAGCTTCTGCAGTGGCCGTGAGGACCAGCAGTGAGGACCAGCATGAAGCCCACAGGGA  9339

Qy  1893  CTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTTGTCCAGGGTCCCTGGGAGAGAGC  1952
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9340  CTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTCTTGTCCAGGGTCCCTGGGAGAGAGC  9399

Qy  1953  CATCTCGCCAAACAAAAGTGCCCACTATATCAACCACGAGACTCAAACA  2001
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  9400  CATCTCGCCAAACAAAAGTGCCCACTATATCAACCACGAGACTCAAACA  9448
```

; RESULT 15
; US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-22

Query Match   59.1%;   Score 1182.6;   DB 9; Length 13957;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy   813  GGTACCTACTCATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTG   872
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8260  GGAAGAAACTCATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTG  8319

Qy   873  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGTCTACAGGATGCTAAGGCT   932
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8320  GCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGTCTACAGGATGCTAAGGCT  8379

Qy   933  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA   992
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8380  CCTAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGA  8439

Qy   993  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG  1052
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8440  AATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAG  8499

Qy  1053  ATCCCTGGAAGTTCCGATGATGCAGTCCGTTTACAAAGACGTTTGGATAACATGAACTT  1112
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8500  ATCCCTGGAAGTTCCGATGATGCAGTCCGTTTACAAAGACGTTTGGATAACATGAACTT  8559
```

```
Qy   1113  CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGTTGGAAGCCAGTTC  1172
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8560  CAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGTTGGAAGCCAGTTC  8619

Qy   1173  TGACCAGTGGAAGGCGTCTGCACCTTTCTCTGCAGGAACTTCGTGTGGCTACAGCTGAA  1232
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8620  TGACCAGTGGAAGGCGTCTGCACCTTTCTCTGCAGGAACTTCGTGTGGCTACAGCTGAA  8679

Qy   1233  AGATGATGAATTAAGCCGGCAGACCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA  1292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8680  AGATGATGAATTAAGCCGGCAGACCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCA  8739

Qy   1293  GAACGATGTACATAGGGCCTTCAAGAGAGGAATTGAAAACTAAAGAACCTGTAATCATGAG  1352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8740  GAACGATGTACATAGGGCCTTCAAGAGAGGAATTGAAAACTAAAGAACCTGTAATCATGAG  8799

Qy   1353  TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT  1412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8800  TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACT  8859

Qy   1413  CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCT  1472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8860  CTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCT  8919

Qy   1473  ACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGAAAAATTGAACCTGCACTCCGCTGA  1532
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8920  ACGAAAGCACGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA  8979

Qy   1533  CTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTTCAGAGACTTCAAGAGGCCACGGA  1592
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8980  CTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTTCCAGGACTTCAAGAGGCCACGGA  9039

Qy   1593  TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG  1652
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9040  TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGG  9099

Qy   1653  CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGCACTTCGAGGAGA  1712
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9100  CGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGCACTTCGAGGAGA  9159

Qy   1713  AATTGCGCGCCCTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCCAGCTTACCAC  1772
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9160  AATTGCGCGCCCTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCCAGCTTACCAC  9219

Qy   1773  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG  1832
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9220  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATG  9279

Qy   1833  GAAGCTTCTGCAGGTGGCCGTCGAGGACCCGGATCAGGCAGCTGCATGAAGCCCACAGGGGA  1892
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9280  GAAGCTTCTGCAGGTGGCCGTCGAGGACCCGGATCAGGCAGCTGCATGAAGCCCACAGGGGA  9339

Qy   1893  CTTTGGTCCAGCATCCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC  1952
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9340  CTTTGGTCCAGCATCCAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGC  9399

Qy   1953  CATCTCGCCAAACAAAGTGCCCTACTACTATCACCACGAGACTCAAACA  2001
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   9400  CATCTCGCCAAACAAAGTGCCCTACTATCATCACCACGAGACTCAAACA  9448
```

Search completed: April   5, 2004, 16:33:02
Job time : 498.207 secs

OM nucleic - nucleic search, using sw model

```
             Copyright   GenCore version 5.1.6
                    (c) 1993 - 2004 Compugen Ltd.
```

Run on: April 4, 2004, 16:52:43 ; Search time 3498.53 Seconds
                   (without alignments)
                   17079.796 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcattgatgggagag.......tcaaccacgagactcaaaca 2001

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : EST:*
            1: em_estba:*
            2: em_esthum:*
            3: em_estin:*
            4: em_estmu:*
            5: em_estov:*
            6: em_estpl:*
            7: em_estro:*
            8: em_htc:*
            9: em_estl:*
           10: gb_est2:*
           11: gb_htc:*
           12: gb_est3:*
           13: gb_est4:*
           14: gb_est5:*
           15: em_estfun:*
           16: em_estom:*
           17: em_gss_hum:*
           18: em_gss_inv:*
           19: em_gss_pln:*
           20: em_gss_vrt:*
           21: em_gss_fun:*
           22: em_gss_mam:*
           23: em_gss_mus:*
           24: em_gss_pro:*
           25: em_gss_rod:*
           26: em_gss_phg:*
           27: em_gss_vrl:*
           28: gb_gss1:*
           29: gb_gss2:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| Result No. | Score | % Match | Query Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 678.4 | 33.9 | 3870 | 11 | BC036103 | BC036103 Homo sapi |
| 2 | 678.4 | 33.9 | 5691 | 29 | AY399453 | AY399453 Homo sapi |
| 3 | 591.8 | 29.6 | 834 | 12 | BI729851 | BI729851 Mus muscu |
| 4 | 547.4 | 27.4 | 3056 | 11 | AK044536 | AK044536 Mus muscu |

| No. | Score | % | Length | DB | ID | Accession | Organism |
|---|---|---|---|---|---|---|---|
| 5 | 547.4 | 27.4 | 5697 | 29 | AY399455 | AY399455 | Mus muscu |
| 6 | 491.4 | 24.6 | 750 | 12 | BI730168 | BI730168 | 603349711 |
| 7 | 486.6 | 24.3 | 579 | 9 | AL121550 | AL121550 | DKFZp762L |
| 8 | 471.6 | 23.6 | 728 | 14 | CB228986 | CB228986 | AGENCOURT |
| 9 | 459.4 | 23.0 | 1047 | 14 | CB850319 | CB850319 | MRA-0070 |
| 10 | 446 | 22.3 | 5676 | 29 | AY399454 | AY399454 | Pan trogl |
| 11 | 424.2 | 21.2 | 663 | 12 | BM488464 | BM488464 | pgm2n.pk0 |
| 12 c | 409 | 20.4 | 595 | 14 | CB177816 | CB177816 | is21c01.x |
| 13 | 389 | 19.4 | 644 | 13 | BU313510 | BU313510 | 603540290 |
| 14 c | 324.6 | 16.2 | 402 | 14 | CB547284 | CB547284 | AMGNNUC:S |
| 15 | 310.8 | 15.5 | 2874 | 29 | AY408546 | AY408546 | Homo sapi |
| 16 | 309.8 | 15.5 | 3753 | 11 | AK081426 | AK081426 | Mus muscu |
| 17 | 301.2 | 15.1 | 2874 | 29 | AY408548 | AY408548 | Mus muscu |
| 18 | 278.2 | 13.9 | 9915 | 29 | AY407022 | AY407022 | Homo sapi |
| 19 | 265.4 | 13.3 | 1122 | 14 | CF109978 | CF109978 | Shultzoni |
| 20 | 265.4 | 13.3 | 3051 | 11 | BC036095 | BC036095 | Homo sapi |
| 21 | 261.2 | 13.1 | 2334 | 11 | BC011062 | BC011062 | Mus muscu |
| 22 | 261.2 | 13.1 | 9691 | 29 | AY407024 | AY407024 | Mus muscu |
| 23 | 257.8 | 12.9 | 1541 | 11 | AK034383 | AK034383 | Mus muscu |
| 24 | 255.4 | 12.8 | 9772 | 29 | AY407023 | AY407023 | Pan trogl |
| 25 | 252.4 | 12.6 | 250 | 29 | BI250598 | BI250598 | 602993659 |
| 26 | 233.6 | 11.7 | 826 | 14 | CD656896 | CD656896 | AGENCOURT |
| 27 | 233 | 11.6 | 835 | 12 | BI553820 | BI553820 | 601190772 |
| 28 | 230.8 | 11.5 | 784 | 12 | BG212445 | BG212445 | RST32032 |
| 29 | 230.2 | 11.5 | 763 | 14 | CB518960 | CB518960 | UI-M-GH0- |
| 30 | 229.2 | 11.5 | 778 | 14 | CB524596 | CB524596 | UI-M-FY0- |
| 31 | 226 | 11.3 | 250 | 14 | BF963618 | BF963618 | QV2-NN004 |
| 32 | 222.2 | 11.1 | 2874 | 29 | AY408547 | AY408547 | Pan trogl |
| 33 | 216.6 | 10.8 | 681 | 13 | BU301653 | BU301653 | 603609005 |
| 34 | 199.8 | 10.0 | 696 | 13 | BJ075057 | BJ075057 | BJ075057 |
| 35 | 199.2 | 10.0 | 697 | 13 | BU456556 | BU456556 | 603777141 |
| 36 | 184.4 | 9.2 | 812 | 14 | CF748008 | CF748008 | UI-M-HE0- |
| 37 c | 182.4 | 9.1 | 645 | 14 | AW467977 | AW467977 | he29g08.x |
| 38 | 179.6 | 9.0 | 466 | 14 | CD549993 | CD549993 | B0305E01- |
| 39 | 172.2 | 8.6 | 630 | 9 | AL855376 | AL855376 | AL855376 |
| 40 | 167.2 | 8.4 | 502 | 9 | AL602076 | AL602076 | DKFZp313B |
| 41 c | 167.2 | 8.4 | 646 | 12 | BI289102 | BI289102 | UI-R-DK0- |
| 42 | 165.4 | 8.3 | 851 | 13 | BU201022 | BU201022 | 603952191 |
| 43 | 164.6 | 8.2 | 600 | 13 | BI988528 | BI988528 | 4012-24 M |
| 44 c | 164.2 | 7.9 | 569 | 29 | CE228097 | CE228097 | tigr-gss- |
| 45 | 158.2 | 7.9 | 405 | 14 | CB811308 | CB811308 | AMGNNUC:S |

ALIGNMENTS

RESULT 1
BC036103

```
LOCUS       BC036103                3870 bp    mRNA    linear   HTC 19-NOV-2003
DEFINITION  Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker
            types), mRNA (cDNA clone IMAGE:5274415), with apparent retained
            intron.
ACCESSION   BC036103
VERSION     BC036103.1  GI:23271310
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3870)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
```

```
Qy    301  TGGGAATGCCTCAGGTAGTGACATGGAAAAAACAAAGCAATTTACATAGAGTTTTAATG  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1443  TGGGAATGCCTCAGGTAGTGACATGGAAAAAACAAAGCAATTTACATAGAGTTTTAATG 1502

Qy    361  GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAAGA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1503  GATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAAAGA 1562

Qy    421  ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTA  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1563  ACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTA 1622

Qy    481  CAACAACATAAGGTGCTTCAAGAAGATCAGAACAAGCAAGTCAGGGTCAATTCTCTC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1623  CAACAACATAAGGTGCTTCAAGAAGATCAGAACAAGCAAGTCAGGGTCAATTCTCTC 1682

Qy    541  ACTCCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCCGCAACTGCTGCTTTGGAA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1683  ACTCCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCCGCAACTGCTGCTTTGGAA 1742

Qy    601  GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTTGTAGTGACAGAAGACCGC  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1743  GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTTGTAGTGACAGAAGACCGC 1802

Qy    661  TGGGTTCTTTACAAGACCACCAGCCTGACCTAGCTCCTGGACTGAC  704
           ||||||||||||||||| | |||||||||||||||||||||||||
Db   1803  TGGGTTCTTTACAAGACCATCCTTCTCAAATGGCAACGTCTTAC 1846
```

RESULT 2
AY399453
LOCUS   AY399453   5691 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION   Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
               genomic survey sequence.
ACCESSION   AY399453
VERSION   AY399453.1 GI:39755442
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 5691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
               Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
               Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
               Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
               gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
  PUBMED   14671302
REFERENCE   2 (bases 1 to 5691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
               Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
               Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
               Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
               Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
               them based on alignment.
FEATURES   Location/Qualifiers
     source   1..5691
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>5691
                     /locus_tag="HCM0229"
ORIGIN

Query Match   33.9%;   Score 678.4; DB 29; Length 5691;
Best Local Similarity   97.7%; Pred. No. 2.9e-142;
Matches 688; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

               Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
               Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
               Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
               Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
               Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
               human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
  PUBMED   12477932
REFERENCE   2 (bases 1 to 3870)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (31-JUL-2002) National Institutes of Health, Mammalian
               Gene Collection (MGC), Cancer Genomics Office, National Cancer
               Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
               USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
               cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
               Toshiyuki and Piero Carninci (RIKEN)
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Sequencing Group at the Stanford Human Genome
               Cencer, Stanford University School of Medicine, Stanford, CA 94305
               Web site: http://www-shgc.stanford.edu
               Contact: (Dickson,Mark) mcd@paxil.stanford.edu
               Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
               R. M.

               Clone distribution: MGC clone distribution information can be found
               through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
               Series: IRAK Plate: 48 Row: f Column: 7
               This clone was selected for full length sequencing because it
               passed the following selection criteria: Similarity but not
               identity to protein
               This clone has the following problem: retained intron.
FEATURES   Location/Qualifiers
     source   1..3870
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5274415"
                     /tissue_type="Brain, hippocampus"
                     /clone_lib="NIH_MGC_95"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
ORIGIN

Query Match   33.9%;   Score 678.4; DB 11; Length 3870;
Best Local Similarity   97.7%; Pred. No. 2.7e-142;
Matches 688; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1  GGCAGTTCATTGATGGAGGAGAGTAAACCTGGACCGTTATCAAACAGCTTTAGGAAGAA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1143  GGCAGTTCATTGATGGAGGAGAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAA 1202

Qy     61  GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGGACAACAAGGAGAGATTCTAAT  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1203  GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGGACAACAAGGAGAGATTCTAAT 1262

Qy    121  GATGTGGAAGTGGTGAAAGACCAGTTCATATCATGAGGGGTACATGATGGATTTGACA  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1263  GATGTGGAAGTGGTGAAAGACCAGTTCATATCATGAGGGGTACATGATGGATTTGACA 1322

Qy    181  GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGACAAGTAAGCTAAGCAGAGGA  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1323  GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGACAAGTAAGCTAAGCAGAGGA 1382

Qy    241  AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATCTCCTAAATTCAAGA  300
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   1383  AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 1442
```

```
Qy     1 GGCAGTTCATTGATGGAGGAGTGAAGTAAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   385 GGCAGTTCATTGATGGAGGAGTGAAGTAAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 444

Qy    61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCTAAT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   445 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACACAGGAGAGATTTCTAAT 504

Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGGATTTGACA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   505 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGGATTTGACA 564

Qy   181 GCCCATCAGGGCCGGGTTGGTATATTCTACAATTGGAAGTGATTGGAACAGGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   565 GCCCATCAGGGCCGGGTTGGTATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGA 624

Qy   241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCACAGAGCAGATGAATCTCCTAAATTCAAGA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCACAGAGCAGATGAATCTCCTAAATTCAAGA 684

Qy   301 TGGGAATGCCTCAGGGTAGCTAGCTGCATGGAAATCAGAAAACAAGCAATTTACATAGAGTTTTAATG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   685 TGGGAATGCCTCAGGGTAGCTAGCTGCATGGAAATCAGAAAACAAGCAATTTACATAGAGTTTTAATG 744

Qy   361 GATCTCCAGAATCAGAACTGAAAGAGTTGAATGAGTTGGCTAACAAAACAGAAGAAGAAGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   745 GATCTCCAGAATCAGAACTGAAAGAGTTGAATGAGTTGGCTAACAAAACAGAAGAAGAAGA 804

Qy   421 ACAAGGAAAATGGAGGAAGAGCCTCTTGAGACCTGATCTTGAAGACGCCAAGTA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805 ACAAGGAAAATGGAGGAAGAGCCTCTTGAGACCTGATCTTGAAGACGCCAAGTA 864

Qy   481 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   865 CAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 924

Qy   541 ACTCACACATGGTGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   925 ACTCACACATGGTGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 984

Qy   601 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   985 GAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGC 1044

Qy   661 TGGGTTCTTTTACAAGACCAGCCTAGCACCAGCCTCCTGGACTGAC 704
         ||||| |||| || |||||||| |||||||||||| |
Db  1045 TGGGTTCTTTTACAAGACATCCTTTCAAATGGCAACGTCTTAC 1088
```

RESULT 3
BI729851
LOCUS   BI729851   834 bp   mRNA   linear   EST 20-SEP-2001
DEFINITION 603349511F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
            mRNA sequence.
ACCESSION   BI729851
VERSION   BI729851.1   GI:15706864
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 834)
AUTHORS   The Cepko Laboratory.
TITLE   Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
JOURNAL   Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1908 row: e column: 03
            High quality sequence stop: 796.
FEATURES   Location/Qualifiers
     source   1..834
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5357162"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_94"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

```
Query Match             29.6%;    Score 591.8;   DB 12;   Length 834;
Best Local Similarity   90.5%;    Pred. No. 7, 1e-123;
Matches   665;  Conservative    0;  Mismatches    67;   Indels    3;   Gaps    3;

Qy   813 GGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTG 872
         |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    75 GGAGGAAACTCATAGATTACTGCAGCAGTTCCCTCTGGACCTGGAGAGTTTCTTCTTG 134

Qy   873 GCTTACAGGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCT 932
         ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   135 GATTACGGAGCAAGCAGAAACAAC-GCCAATGTCCTACAGGACGCTTCCCGTAAGGAGAAGAGCT 194

Qy   933 CCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACACTTCCAAGGTGA 992
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   195 CCTAGAAGACTCC-AGGGAGTCAGAGAGCTCAGAGAGCTGATGAAACCATGGCAAGACTCTCCAACAGGAGA 253

Qy   993 AATTGAAGCTCACACAGATGTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAG 1052
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 AATTGAAACTCACACAGATATCATCACAATCTTGATGAAAATGGCCAAAAAATCCTGAG 313

Qy  1053 ATCCCTGGAAGGTTCCGATGATCGCAGTCGTCCTGTTACAAAGACGTTTGATACACTGAACTT 1112
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314 ATCCCTGGAAGGTTCGGATGAAGCACCCCGTTACAAAGACGTTTGATAAACATGAATTT 373

Qy  1113 CAAGTGGGAGTGAACTTCGGGAAAAAGTCTCAACATTAGTCCCATTGGAAGCCAGTTC 1172
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   374 CAAGTGGGAGTGAACTTCAGGAAAAAGTCTCAACATTAGGTCCCATTGGAAGCAAGTTC 433

Qy  1173 TGACCAGTGGAAGCGTCTGCAACCGTCCTCGGTGTGGCTACCAGCTGAA 1232
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   434 TGACCAGTGGAAGCGTTGCATCTTTCTTGCATCTTCTTCTTCTTCCAGGACTTCTTGTTTGGCTACCAGCTGAA 493

Qy  1233 AGATGATGAATAAGCCGGCAGCCATTGGAGGCGCACTTTCCAGCAGTTCAGAAGCA 1292
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   494 AGATGATGAACTGAACGCCGTCAGGCACCATCCGTGTGATTTCCCAGCAGTTCAGAAGCA 553

Qy  1293 GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 1352
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   554 GAATGATATACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAG 613

Qy  1353 TACTCTTGGACTGTGACTAGGCGTCAATATTTCTGACAGAGGTCAGCCTTTGGAAGGACTAGAGAAACT 1412
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   614 TACTCTGGAGACTGTGAGATATTTCTGACAGAGGTCAGCCTTTGGAAGGACTAGAGAAACT 673

Qy  1413 C-TACCAGGAGCCCAGAGAGCTGCCTCCTGGAGAGGAATGCCAGAATGCTCAGAATGTCACTCGGCTTC 1471
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   674 CTTACCAGGAGCCCAGAGAGCTGCCTCCTCCTGGAGAAGAAGAGCTCGAAAGTCACTCGGCTCC 733

Qy  1472 TACGAAAGCAGCAGGTCTACGAATATTTCTGAGGTCAATACTGAGTGGGAAAAATTGAACCTGACC 1530
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   734 TACGAAAGCAGGTCGAAGAGGTCAACCCTGAATTGGGACCAATTGAACCTGCCTCAAGCT 793

Qy  1531 GACTGGCAGAGAAA 1545
```

```
Db         794 GATTGGCAGAGAAAA 808
               ||||||||||||||||

RESULT 4
AK044536
LOCUS       AK044536                3056 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus adult retina cDNA, RIKEN full-length enriched library,
            clone:A930019F21 product:dystrophin, muscular dystrophy, full
            insert sequence.
ACCESSION   AK044536
VERSION     AK044536.1  GI:26090404
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS     Carninci,P. and Hayashizaki,Y.
TITLE       High-efficiency full-length cDNA cloning
JOURNAL     Meth. Enzymol. 303, 19-44 (1999)
MEDLINE     99279253
PUBMED      10349636
REFERENCE   2
AUTHORS     Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE       Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL     Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE     20499374
PUBMED      11042159
REFERENCE   3
AUTHORS     Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazana,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE       RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
JOURNAL     Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE     20530913
PUBMED      11076861
REFERENCE   4
AUTHORS     The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
TITLE       Functional annotation of a full-length mouse cDNA collection
JOURNAL     Nature 409, 685-690 (2001)
REFERENCE   5
AUTHORS     The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
TITLE       Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL     Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3056)
AUTHORS     Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,J., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
TITLE       Direct Submission
JOURNAL     Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group; RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Retina RNA was provided by Dr. Stefano Gustincich (Department of
            Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
            02115, USA) whose assistance is gratefully acknowledged. Please
            visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..3056
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:A930019F21"
                     /db_xref="MGI:2409854"
                     /db_xref="taxon:10090"
                     /clone="A930019F21"
                     /tissue_type="retina"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    1..3056
                     /note="dystrophin, muscular dystrophy (MGD|MGI:94909,
                     GB|NM_007868, evidence: BLASTN, 100%, match=999)"
ORIGIN

Query Match            27.4%;    Score 547.4;  DB 11;   Length 3056;
Best Local Similarity  88.0%;    Pred. No. 1e-112;
Matches  596;    Conservative   0;    Mismatches   81;   Indels    0;   Gaps    0;

Qy      1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTTAGGAGAA 60
          |||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||
Db   1044 GACAGTTCATTGATGGAGACGGAGAGTAAAATCTGGATAGTTACCAAACTGCTTTAGGAAGAA 1103

Qy     61 GTATTATCGTGGCTTCTTCTGCTGGAGGACACATTGCAAGGACACAGGAGAGATTTCTAAT 120
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   1104 GTACTTTCATGGCTTCTTCTGCCGAGGATACATTGCGAGGCACAAGGACAGGAGAGATTTCAAAT 1163

Qy    121 GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGATTGACA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1164 GATGTTGAAGAAGTGAAAGAACAGTTTCATGCTCATGAGGGATCATGAGGATCTGACA 1223

Qy    181 GCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGA 240
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1224 TCTCATCCAAGGACTTGTTGGTAATGTTCACAGTTAGGAGTCAACTAGTTGGAAAAGGG 1283

Qy    241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGCAGCAGATGAATCTCCTAAATTCAAGA 300
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1284 AAATTATCAGAAGATGAAGAAGAGCTGAAGTGCAAGAACAAATGAATCTCCTAAATTCAAGA 1343

Qy    301 TGGGAATGCCTCAGGTAGCTAGACATGGAAAAACACAAGCAATTTACATAGAGTTTTATG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1344 TGGGAATGTCTCAGGTAGCTAGCATGGAAAAACAAAGCAAAACAAAGCAAAATTACACAAGTTCTAATG 1403

Qy    361 GATCTCCAGAATCAGAAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 420
          ||||||||||||||||||| |||||||||||||||||| |||||||
Db   1404 GATCTCCAGAATCAGAAATTAAAAGAAACTAGAGACTGCTTAACAAAAACTGAAGAAGAGA 1463

Qy    421 ACAAGGAAAATGGAGGAAGAGCCCTTCAAGACCTTTGAAGACCTAAAACGCCAAGTA 480
          ||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1464 ACTAAGAAAATGGAGGAAGAGCCCCTTTGGACCTTTGAAGATCTTGAAGAATGCCAAGTA 1523

Qy    481 CAACAACATAAGGTGCTTCAAGAAGATCTAGAACCAAGACAAGTCAGGGTCAATTCTCTC 540
          |||||||||||||| |||||||||||||||||||| ||||||| ||||||||||||||||
Db   1524 CAACAACATAAGGTGCTTCAAGAAGATCTAGAAGATCTAGAAGAACAGGACAGGCAGGGTCAGGTCAACTCGCTC 1583

Qy    541 ACTCACATGGTGGTGTGGAGTTGATGAATCTAGTGATGAATCCAGTGATCACGCAACTGCTGCTTGGAA 600
          |||||||| || |||||||  |||||  ||| |||||||||||||||||||||
```

Left column:

```
Db   1584 ACTCACATGGTAGTGGTTGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTGGAA 1643
Qy    601 GAACAACTTAAGGTAGTATTGGGAGATCGATGGCAAACATCGTAGATGGACAGAAGACCGC 660
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1644 GAACAACTTAAGGTACTGGGAGATCGATGGGCAAATATCGCAGATGGACTGAAGAACCGC 1703

Qy    661 TGGGTTCTTTTACAGAA 677
          ||||| || |||||||
Db   1704 TGGATTGTTTTACAGAA 1720

RESULT 5
AY399455
LOCUS       AY399455           5697 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION  Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY399455
VERSION     AY399455.1   GI:39755444
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..5697
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>5697
                     /locus_tag="HCM0229"
ORIGIN

Query Match    27.4%;   Score 547.4;  DB 29;  Length 5697;
Best Local Similarity  88.0%;  Pred. No. 1.2e-112;
Matches  596;  Conservative  0;  Mismatches  81;  Indels  0;  Gaps  0;

Qy     1 GGCAGTTCATTGATGGAAGTGAAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAA 60
         | |||||||||| |||||||||| |  ||||   |||||| || |||| ||||||||||||
Db   391 GACAGTTCATTGATGGATGAGACGACGGAAGTAAATCTGATAGTTACCAAACTGCTTTAGAGAGA 450

Qy    61 GTATTATCGTGGCTTCTTTCTGCGTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAT 120
         | |||||||||||| ||  |  || |||| ||| |||||||| ||| ||||||| |  ||
Db   451 GTACTTTCATGGCTTCTTTCTGCGGAGGATAACATTGCCAGCACCAAGGAGAGATTTCAAAT 510

Qy   121 GATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGGATTTGACA 180
         |||| |||||  |||||||| |||| || |||||||| |   |||||||||||||||
Db   511 GATGTTGAAGAAGTGAAAGAACAGTTCATGCTTCATGAGGGATTCATGATGATGGATCTGACA 570

Qy   181 GCCCATCAGGCGCCGGTTGGTAATATTCTACAATTGGAAGTGAAGCTGATTGGAACAGGA 240
         |||| ||| ||  || || ||||||| |||| ||| |||| |||| |  | |||||||
Db   571 TCTCATCAAGGACTTGTTGGTAATGTTCTACAGTTCAACTAGTTGGAAGAGTCAACTAGTTGGAAAAGGG 630

Qy   241 AAATTATCAGAAGATGAAGAAACTCGAAGTACAAGAGCAGCAGATGAATCTTCCTAAATTCAAGA 300
         ||||||| |||
```

Right column:

```
Db   631 AAATTATCAGAAGATGAAGAAGCTGAAGTGCAAGAACAAACAAATGAATCTCCTAAATTCAAGA 690

Qy   301 TGGGAATGCCTCCAGGGTAGTCAGCATGGAGAAAAACAAAGCAATTCACATAGAGTTTTAATG 360
         ||||||||| || |||||||||||||||||| |||||| |||||||| ||||||||||
Db   691 TGGGAATGTTCAGGGTAGCTAGCATGGAGAAAAACAAAGCAAATACACAAAGTTCTAATG 750

Qy   361 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTAACAGAAAACAGAGAAGAAA 420
         ||||||||||||||||||  |||||   ||||||||||||| ||  ||||||||||||
Db   751 GATCTCCAGAATCAGAAATTAAAAGAACTAGATGACTGGTTAACAAAAAACTGAGAGAGAGA 810

Qy   421 ACAAGGAAAATGGAGGAAGAGCCCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTA 480
         |||| ||||||||||||||||||| | ||||| |||||||||||| |||||| ||||||
Db   811 ACTAAGAAAATGGAGGAAGAGCCCTTGGACCTTGAAGATCTAAAATGCCAAGTA 870

Qy   481 CAACACATAAGGTGCTTCAAGAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTC 540
         |||||||||||||||| ||||||||| ||||||||| |||||||||||||||||||||
Db   871 CAACAACATAAGGTGCTTCAAGAAGATCTAGAAGAAGATCAACAGGAGCAGGGTCAACTCGCTC 930

Qy   541 ACTCACATGGTGGTGTAGTTGATGAATTAGTGGAGATCACGCAACTGCTGCTTTGAA 600
         |||||||||||||||| |||||||| |||||||||||| ||||||||||||||||
Db   931 ACTCACATGGTAGTGTGGTTGATGAATCCAGCGGTCGTGATCATGCAACAGCTGCTTTGGAA 990

Qy   601 GAACAACTTAAGGTGGTATTGGGAGATCGATGGCAAACATCTGTAGATGGACAGAAGACCGC 660
         |||||||||||||| |||||||||||||||||||| |||| ||||||||||||||||||||
Db   991 GAACAACTTAAGGTACTGGGAGATCGATGGGCAAATATCTGCAGATGGACTGAAGAACCGC 1050

Qy   661 TGGGTTCTTTTACAGA 677
         ||||| |||||||||
Db  1051 TGGGATTGTTTTACAGA 1067

RESULT 6
BI730168
LOCUS       BI730168  60349711F1 NIH_MGC_94  Mus musculus cDNA clone IMAGE:5357187 5',
            750 bp  mRNA   linear   EST 20-SEP-2001
DEFINITION  601R9F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357187 5',
            mRNA sequence.
ACCESSION   BI730168
VERSION     BI730168.1   GI:15707181
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 750)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1908  row: f  column: 04
            High quality sequence stop: 747.
FEATURES             Location/Qualifiers
     source          1..750
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5357187"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_94"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN
```

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZp762L078) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.

```
FEATURES             Location/Qualifiers
     source          1..579
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp762L078"
                     /tissue_type="melanoma (MeWo cell line)"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="762 (synonym: hmel2)"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN
```

Query Match
Best Local Similarity   24.3%;   Score 486.6; DB 9;   Length 579;
Conservative 97.2%;   Pred. No. 3.9e-99;
Matches 495;   Mismatches 0;   Indels 14;   Gaps 0;

```
Qy   977 AAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACA 1036
             |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    71 AGGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACA  130

Qy  1037 GCCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCGTGTTACAAAGACGTT 1096
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 GCCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCGTGTTACAAAGACGTT  190

Qy  1097 TGGATAACATGAACTTCAAGTGGAGTGGAACTTCAGAAAAAGTCTCTCAACATTAGGTCCC 1156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 TGGATAACATGAACTTCAAGTGGAGTGGAACTTCAGAAAAAGTCTCTCAACATTAGGTCCC  250

Qy  1157 ATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGG 1216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 ATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGG  310

Qy  1217 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGGGGACTTTC 1276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGGGGACTTTC  370

Qy  1277 CAGCAGTTCAGAAGACAGAAGCAGAAGATGTACATAGGGAATTGAAAACTAAAG 1336
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 CAGCAGTTCAGAAGACAGAAGCAGAAGATGTACATAGGGAATTGAAAACTAAAG  430

Qy  1337 AACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGG 1396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 AACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGG  490

Qy  1397 AAGGACTAGAGAAACTCTACCAGAGCCCAGAGAGCTGCCTCCTGGAGGAGAGCCCAGA 1456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 AAGGGCTAGAGAAACTCTACCAGAGCCCAGAGAGCTGCCTCCTGGAGGAGAGCCCCAG  550

Qy  1457 ATGTCACTCGGCTTCACGAAAGCAGGCT 1485
             ||||||||||||||||||||||||||||
Db   551 AATGTCACTCCGCTTCACGAAAGCAGGCT  579
```

RESULT 8
CB228986
```
LOCUS       CB228986                728 bp    mRNA       linear   EST 10-FEB-2003
DEFINITION  AGENCOURT 11499247 NICHD_Rh_Ov1 Macaca mulatta cDNA clone
            IMAGE:6884820 5', mRNA sequence.
ACCESSION   CB228986
VERSION     CB228986.1  GI:28280564
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
```

Query Match
Best Local Similarity   24.6%;   Score 491.4; DB 12;   Length 750;
Conservative 88.4%;   Pred. No. 3.4e-100;
Matches 602;   Mismatches 0;   Indels 8;   Gaps 6;

```
Qy   804 GATGTTGGAGGTACCACTCACTAGATTACTGCAACAGTTCCCCTGGACCTGGAAAGTT 863
             | |  ||| ||||| |||||||||||||||| ||||||||||||||||||||||| ||
Db    69 GCTGCTTGGGACAGAAACTCATAGATTACTGCAGCAGTTCCCTGGACCTGGAGAAGTT 128

Qy   864 TCTTGCCTGG-CTTACAGAAGCTGAAAC--AACTGCCAATGTCCTACAGGATGCTACCCG 920
             ||||  ||||  ||||| ||||| ||||  | ||| |||||||||||||||| | |||||
Db   129 TCTTTCCTGGATTTACGGAAGCAGAGAACACAGAGCTGCCAATGTCCTACAGGACGCTTCCCG 188

Qy   921 -TAAGGAAAGGCTCCTAGAGACTCCAAGG--AGTAAAAGAGCTGATGAAACAATGGCA 977
              ||||||| |||||||||| |||||||||   |||| | |||||||||||| ||||||
Db   189 CTAAGGAGAAGCTCCTAGCAAGACTCCAGGGGCAGTCAGGGGCAGTGCAGAGAGCTGATGAAACCATGGCA 248

Qy   978 AGACCTCCAAGTGAAA-TTGAAGCTCACACAGATGTTTATCACAAACCTGGATGAAAACA 1036
             |||| |||||| ||||  |||||| ||||||||| | |||||||  | |||||||||| |
Db   249 AGATCTCCAAGAGAAGAAAGTTGAAACTCACACAGATATCTATCACACAATCTTGATGAAAATG 308

Qy  1037 GCCAAAAAATCCTGGAAGTCCGATGATGCAGTCCGTGTTACAAAGACGTT 1096
             |||||||||||| ||| ||||||||||||| |||||||||||||||||||
Db   309 GCCAAAAAATCCTGGAGTCCGATGATGCAGTCCGTGTTACAAAGACGTT 368

Qy  1097 TGGATAACATGAACTTCAAGTGGAGTGGAACTTCAGAAAAAGTCTCTCAACATTAGGTCCC 1156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   369 TGGATAACATGAACTTCAAGTGGAGTGGAACTTCAGAAAAAGTCTCTCAACATTAGGTCCC 428

Qy  1157 ATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTTCGCATCTTTCTCTGCAGGAACTTCTTGG 1216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   429 ATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTTCGCATCTTTCTCTGCAGGAACTTCTTG 488

Qy  1217 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGGCGACTTTC 1276
             ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   489 TTTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCACCGTCAGGCGTGGTGATTTCC 548

Qy  1277 CAGCAGTTCAGAAGCA-GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAA 1335
             |||||||||||||| | |||  | ||||||||| |||||||||| ||||||||||||||
Db   549 CAGCAGTTCAGAAGCACGAATGATATACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAA 608

Qy  1336 GAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAGCCCAGAGAGCTGCCTCCTTTG 1395
             ||||||||||||||||||| |||||||||||||||| |  || ||||| ||| ||||| |
Db   609 GAACCTGTAATCATGAGTACTCTGGAGACTGTACTCGAGATATTTCTGACAGAGCAGCCTTTG 668

Qy  1396 GAAGGACTAGAGAAACTCTACCAGAGCCCGCCTCCTGCCTGCTGGAGAGAGCCCCAG 1455
             |||||||||||||||||||||||||||||  ||||| || ||| ||||||| |||||
Db   669 GAAGGACTAGAGAAACTCTACCAGAGCCCAGAGAACTGCCTCCTGGAGGAGAAAGAGCTCAG 728

Qy  1456 AATGTCACTCGGCTTCACGA 1476
             ||||||||||||||||||||
Db   729 AATGTCACTCGGCTTCACGA 749
```

RESULT 7
AL121550
```
LOCUS       AL121550                579 bp    mRNA       linear   EST 04-SEP-2003
DEFINITION  DKFZp762L078_r1 762 (synonym: hmel2) Homo sapiens cDNA clone
            DKFZp762L078 5', mRNA sequence.
ACCESSION   AL121550
VERSION     AL121550.1  GI:5927551
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 579)
AUTHORS     Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
TITLE       EST (Ottenwaelder, et al.)
JOURNAL     Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
```

ORGANISM   Macaca mulatta
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
           Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 728)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Dr. Eliot Spindel
           cDNA Library Preparation: CLONTECH
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM3135 row: e column: 11
           High quality sequence stop: 583.
FEATURES   Location/Qualifiers
     source   1..728
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /db_xref="taxon:9544"
                     /clone="IMAGE:6884820"
                     /tissue_type="Ovary"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_Rh_Ov1"
                     /note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
                     Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.0-4.0 Kb. Tissue pooled from
                     pre-pubertal, post pubertal sn menopausal monkeys.
                     Constructed by Clontech. Note: this is a NICHD Library."
ORIGIN

Query Match   23.6%; Score 471.6; DB 14; Length 728;
Best Local Similarity   95.3%; Pred. No. 9.9e-96;
Matches 486; Conservative   0; Mismatches 24; Indels   0; Gaps   0;

Qy   195 GGTTGGTAATATTCTACAATTGGAGTAAGTGAGTGATTGGAACAGGAAAATTATCAGAAGA 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGTTGGTAATATTCTACAATTGGAGTAAGTGAGTGATTGGAACAGGAAAATTATCAGAAGA 60

Qy   255 TGAAGAAACTGAAGTACAAGAGCAGATGAATGCTCCTAAATTCAAGATGGGAATGCCTCAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGAAGAAACTGAAGTACAAGAGCAGATGAATGCTCCTAAATTCAAGATGGGAATGCCTCAG 120

Qy   315 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCA 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCTCCAGAATCA 180

Qy   375 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAGAACAAGGAAAATGGA 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GAAACTGAAAGAGTTGAATGACTGGCTGACAAAAACAGAAGAAGAACAAGGAAAATGGA 240

Qy   435 GGAAGAGGCCTCTTGGACCTGATCTTGAAGACCAAGACGCCAAGTACAACAACAACATAAGGT 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAAAGAGACCCCTTGGACCTTGAGACCTGATCTTGAGACCAAGACGCCAAGTACAACAACAACAACATAAGGT 300

Qy   495 GCTTCAAGAAGATCTAGAACAGAACAAGAACAAGTCAGGGTCAATCTCACTCAGGTGTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCTTCAAGAAGATCTAGAACAGAACAAGAACAAGTCAGGGTCAATCTCACTCCACATGGTGGT 360

Qy   555 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGT 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAGCAACTTAAGGT 420

Qy   615 ATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACA 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ATTGGGAGATCGATGGGCAAACATCTGCAGATGGACAGAAGACCGCTGGGTTCTTTTACA 480

Qy   675 AGACCAGCCAGCTCCTGGACTGAC 704

Db   481 AGACATCCTTCTCAAATGGCAACGTCTTAC 510

RESULT 9
CB850319
LOCUS   CB850319   MRA-0070 MOUSE ADULT RETINA   1047 bp   mRNA   linear   EST 01-SEP-2003
DEFINITION MRA-0070 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA sequence.
ACCESSION   CB850319
VERSION   CB850319.1 GI:34380806
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1047)
AUTHORS   Yu,J., Farjo,R., MacNee,S.P., Baehr,W., Stambolian,D.E. and
            Swaroop,A.
TITLE   Annotation and analysis of 10,000 expressed sequence tags from
            developing mouse eye and adult retina
JOURNAL   Genome Biol. 4 (10), R65 (2003)
MEDLINE   22881944
PUBMED   14519200
COMMENT   Contact: Swaroop, A.
            Department of Ophthalmology and Visual Sciences
            Kellogg Eye Center, University of Michigan
            540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
            Tel: 734 615 2246
            Fax: 734 647 0228
            Email: swaroop@umich.edu.
FEATURES   Location/Qualifiers
     source   1..1047
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="retina"
                     /clone_lib="MOUSE ADULT RETINA"
                     /note="Vector: pSPORT1"
ORIGIN

Query Match   23.0%; Score 459.4; DB 14; Length 1047;
Best Local Similarity   85.8%; Pred. No. 6.1e-93;
Matches 552; Conservative   0; Mismatches 87; Indels   4; Gaps   4;

Qy   33 GGACCGTTATCAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTCTGCTGAGGACAC 92
              ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   86 GGATAGTTACCAACTGCTTTANAAGAGTACTTTCATGTGGAAGTGGTGAAGACCAGTTCATAC 145

Qy   93 ATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTCGTGGAAGAGTGAAGAACAGTTCATAC 152
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   146 ATTGCGAGCACAAGGAGAGATTTCAAATGATCTTGAAGAGTTGAAGAACAGTTTCATGC 205

Qy   153 TCATGAGGGTACATGATGATTGACAGCCCATCACAGGGCCGGGTTGGTATATTCTACA 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206 TCATGAGGGATTCATGATGATGATCTGACATCTCATCAAGGACTTGTTGGTAATGTTCTACA 265

Qy   213 ATTGGGAGTAAGCTAGTGATTGAACAGGAAAATTATCAGAAGATGAAGATGAAGAACTGAAGTACA 272
              ||| ||| ||||||||||||| ||| |||||||||||||||||||||| ||||||||||
Db   266 GTTAGGAAGTCAACTAGTTGAAAAGGGAAAATTATCAGAAGATGAAGATGAAGAAGTGAAGTGCA 325

Qy   273 AGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAG-CTAGCAGCAGTTCATAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   326 AGAACAAATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTANCCTAGCATGCATGGAAA 385

Qy   332 AACAAAGCAATTACATAGAGTTTAATGACTTAAGGT 391
              ||||||||||||||||||||||||||||||||||||
Db   386 AACAAAGCAAATTACACAAAGTTCTAATGATCTCAGAATCAGAAATTAAAAGAACTAG 445

Qy   392 ATGACTGGCTAACAAAAACAGAAGAAGAACAAGAAAATGGAGAAAAGAGCCTCTTGGAC 451
              ||||||||||||| ||||| ||||||| |||||| |||||||||||||||||||| ||||
Db   446 ATGACTGGTTAACAAAAACTGAAGAGAGAACTAANAAAAAATGGAGAGAAGAGCCCTTTGGAC 505

Qy   452 CTGATCTTGAAGACCTAAAAGGCTAAAACGCCAAGTACAACAACAACATAAGGTGC-TTCAAGAAGATCTA 510

```
Db  506 CTGATCTTGGAGATCTAAAATGCCAAGTACAACAACATAAGGTGCTTCAAGAGATCTA 565
Qy  511 GAACAGAAGACAGTCAGGGTCAATTCTTCCACTCACATGGTGGTGGTAGTTGATGAATCT 570
Db  566 GAACAGGAGCAGGTCAGGGTCAACTGCTCTCACTCACGTCGCTCACATGGTGAGTGATGAATCC 625
Qy  571 AGTGGAGATCACGCAACTGCTGTTGGAGAGAACAAC-TTAAGGTATTGGGAGGATCGATG 629
Db  626 AGCGGGGATCATGCAACAGCTGCTTTGGAAANAACAACTTAAGGNACTGGGAGGACTGGGATCGATG 685
Qy  630 GGCAAACATCTG-TAGATGGACAGAGACCCGCTGGGTTCTTTT 671
Db  686 GGCAAATATCTGCAAAAATGGACTGCAAAACCGTTGGATTGTTTT 728
```

RESULT 10
AY399454
LOCUS   AY399454   5676 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY399454
VERSION   AY399454.1 GI:39755443
KEYWORDS   GSS.
SOURCE   Pan troglodytes (chimpanzee)
  ORGANISM Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 5676)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 5676)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..5676
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene   <1..>5676
                     /locus_tag="HCM0229"
ORIGIN

Query Match   22.3%; Score 446; DB 29; Length 5676;
  Best Local Similarity   84.0%; Pred. No. 9.4e-90;
  Matches 458; Conservative   0; Mismatches   87; Indels   0; Gaps   0;

Qy 160 GGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTTGGTAATATTCTACAATTGGGA 219
Db 529 GGGTAGCATGATGGATNNNNNAGCCCATCAGGGCCGGGTTGGTAATANNNNNANTTGGGA 588
Qy 220 AGTAAGCTGATTGGAACAGAGAAAATTATCAGAAGAGAGAAGAAACTGAAGTACAAGAGCAG 279
Db 589 AGTNNNNTGATTGGAACAGAGAAAATTATCAGAAGAGAGAAGAAACTGAAGTACAAGAGCAG 648
Qy 280 ATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGCATGGAAAAACAAAGC 339
Db 649 ATGAATCTNNNNNNNATTCAAGATGGGAANNNNNNNNNNNNNNNNNNNNNNNNNNNN 708
Qy 340 AATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAGAAACTGAAGAGTTGAATGACTGG 399
```

```
Db  709 NNTTTACATAGAGTTTAATGGATCTCCAGAATCAGAAGTGCTTCAAGAAGATCTA 758
Qy  400 CTAACAAAACAGAGAAAGAACAACAGGAAAATGGGAGGAGAGCCTTGGACCTGATCTT 459
Db  769 CTAACAAAACAGAGAAAGAACAACAGGAAAATGGGAGGAGAGCCTTGGACCTGATCTT 828
Qy  460 GAAGACCTAAAAGCGCCAAGTACACAACAACATAAGGTGCTTCAAGAGATCTAGAACAAGAA 519
Db  829 GAAGACCTAAAAGCGCCAAGTACACAACAGTACAACATAAGNNNNNCAAGAAGATTAGAACAAGAA 888
Qy  520 CAAGTCAGGGTCAATTCTTCCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGAT 579
Db  889 CAAGTCAGGGTCAATTCTTCCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGAC 948
Qy  580 CACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGGATCGATGGGCAAACATC 639
Db  949 CACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGGATCGATGGGGCAAACATC 1008
Qy  640 TGTAGATGGACAGAGACCCGTGGGTTCTTTTACAAGACCAGCCTGACCTAGCTCCTGA 699
Db 1009 TGTAGGTGGACAGAAGCCCGCTGGGTTCTTTCTTTTACAGACATCCTTCTCAAATGGCAACCT 1068
Qy  700 CTGAC 704
Db 1069 CTTTAC 1073
```

RESULT 11
BM488464
LOCUS   BM488464   663 bp   mRNA   linear   EST 07-FEB-2002
DEFINITION pgm2n.pk007.117 Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
            clone pgm2n.pk007.117 5' similar to sp|P11533|DMD_CHICK Dystrophin
            pir|S02041 dystrophin, muscle - chicken emb|CAA31746.1| (X11369)
            dystrophin (AA 1 - 3660) [Gallus gallus], mRNA sequence.
ACCESSION   BM488464
VERSION   BM488464.1 GI:18609395
KEYWORDS   EST.
SOURCE   Gallus gallus (chicken)
  ORGANISM Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 663)
  AUTHORS   Cogburn,L.A. and Monsonego-Ornan,E.
  TITLE   ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
            Project
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES   Location/Qualifiers
     source   1..663
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Commercial broiler and Ottawa Res. Centre
                     Strains 90 & 21"
                     /db_xref="taxon:9031"
                     /clone="pgm2n.pk007.117"
                     /sex="Male and Female"
                     /tissue_type="Breast muscle, leg muscle and epiphyseal
                     growth plate"
                     /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
                     ,11 weeks);growth plate(1d,7d,14d post-hatch)"
                     /lab_host="E. coli EMDH10B"
                     /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
                     and Epiphyseal Growth Plate cDNA library (pgm2n)"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
```

ORIGIN

```
        pools of total RNA isolated from each tissue (embryonic
        muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
        plate 33.3% of the final RNA pool). Single pass sequencing
        from 5'-end"
```

Query Match   21.2%;   Score 424.2; DB 12;   Length 663;
Best Local Similarity 78.1%; Pred. No. 5e-85;
Matches   518; Conservative   0; Mismatches 144; Indels   1;   Gaps   1;

```
Qy   960 GCTGATGAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTATCA 1019
             ||||||||||||| | |||||| ||||||||| ||||| ||||||| |||| | |||
Db     1 GCTCATGAAGGAGTGCAGGATCTCAGGCAGGAATTGATGACACATACTGACATCT-TNN 59

Qy  1020 CAACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 1079
             ||||||||||||||||| ||||||||||||| |||||||| ||||||||| | |||
Db    60 CAACCTGGATGAAAACGGGCAGAAAATCCTGAGATCCCTGGAAGGCTCAGGAGGATGCTGT 119

Qy  1080 CCTGTTACAAAGACGTTGGATAACATGAACATTCAAGTGAGTGAACTTCGGAAAAAGTC 1139
             ||||| ||| |||||| ||||| || |||||||||  |||  | ||||| |||| |||
Db   120 CCTGTTGCAGAGACGTCTGATAACATGAACATTCAGATGAGTGAGTGAGTTAGGAAGAAATC 179

Qy  1140 TCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACCTTTC 1199
             ||||| ||||| |||| |||||||||  ||||| ||| ||||||| |||| |||| |
Db   180 TCTAAACATTAGATCTCATTGGAAGCCAGCCAGCACAGACCCAGTGAAGCGTTTACATCTCTC 239

Qy  1200 TCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGAAGATGAATGAAGCCGGCAGGCACC 1259
             ||| |||||||||||| ||||| |||||||| |||||||||||| ||| |||||| |||
Db   240 TCTTCAGGAACTTTTGGCATGGCTGCAATTGCAATTGAAGGAGGATGAATTAAAACAGCAAGCACC 299

Qy  1260 TATTGGAGGGGACTTTCCAGCAGTTCAGAGAGCAGAACGATGTACATAGGGCCTTCAAGAG 1319
             ||| ||||| |||| | |||||||||| ||||| ||||||  | |||||||   |||||
Db   300 CATTGGTGGAGATATTCCCACTGTGCAGAGAGCAGAATGATGTTCATAGGACTTTCAAGAG 359

Qy  1320 GGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCT 1379
             ||||||||| |||| ||||||||| ||||| ||||| || |||||| || |||  | |||
Db   360 GGAGCTGAAAACAAAAGAACCTGTTATCATGAATGCACTTGAGACTGTGCGACTCTTCCT 419

Qy  1380 GACAGAGCAGCCTTTGGAAGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCC 1439
             ||| ||||||| |||||| |||  ||| |||| |||| |||||  |||||||||||| ||
Db   420 GGCAGATCAACCAGTAGAGAGGAGACTCGAAAAGGTCTATCCAGAACCAAGACCTATCACC 479

Qy  1440 TGAGGAGAGCCCAGAAATGTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGGTCAATAC 1499
             |||||||||||||||||||| ||| |||||| |||| |||||| |||||||||||  ||||
Db   480 TGAGGAGGAGGGCCCAGAAATGTCACTAAAGTTCACTAAAGTTCTCCGAAGGCCAAGCAGCAGATGATGTCAGNAC 539

Qy  1500 TGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCT 1559
             |||||||||| ||||||||||| ||||  |||| |||||||||||| ||||||||||| |
Db   540 TGAGTGGGATAAGCTAAATCTCACGTTCTGCTGATTGGCAAAAGAAGAGAATAGATGATGCTCT 599

Qy  1560 TGAAAGACTCCAGGAACTTCAAGAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGGCCAAGC 1619
             ||||||||||||||||||||||| ||  || |||| ||| || || |||||||||||||||
Db   600 TGAAAGAGCTCGAGGGTCTTCAGGAGCATTGATGATGAATCAATGNNNCTGAAACTGCCCAGC 659

Qy  1620 TGA 1622
             |||
Db   660 TGA 662
```

RESULT 12
CB177816/c
LOCUS   CB177816   595 bp   mRNA   linear   EST 31-JAN-2003
DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
            similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ;, mRNA sequence.
ACCESSION   CB177816
VERSION   CB177816.1 GI:28186206
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 595)
AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,

Lenishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.

TITLE   Endocrine Pancreas Consortium
JOURNAL   Unpublished (2000)
COMMENT   Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
          Endocrine Pancreas Consortium
          Harvard University, Howard Hughes Medical Institute
          Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
          MA 02138
          Tel: 617-495-1812
          Fax: 617-495-8557
          Email: dmelton@biohp.harvard.edu
          Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
          Washington University Genome Sequencing Center For information on
          obtaining a clone please contact: Dr. Hiroshi Inoue
          (hinoue@im.wustl.edu)
          Seq primer: -40RP from Gibco
          High quality sequence stop: 448.

FEATURES   Location/Qualifiers
     source   1..595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6553129"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University,
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match   20.4%;   Score 409; DB 14;   Length 595;
Best Local Similarity 100.0%; Pred. No. 1.3e-81;
Matches   409; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 GGCAGTTCATTGATGGAGAGTGAAGTAAACTGGACCGTTATCAAACAGCTTTAGAAGAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   409 GGCAGTTCATTGATGGAGAGTGAAGTAAACTGGACCGTTATCAAACAGCTTTAGAAGAA 350

Qy    61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCACAACAAGGAGAGATTCTAAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   349 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCACAACAAGGAGAGATTCTAAT 290

Qy   121 GATGTGGAAGTGGTAGAAGACCAGTTCATCATGTCAGAAAGTACATGATGGAAGTAAGCTGA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289 GATGTGGAAGTGGTAGAAGACCAGTTCATCATGTCAGAAAGTACATGATGGAAGTAAGCTGA 230

Qy   181 GCCCATCAGGGCCGGTTGTAATATTCACAATTGGAAGTAAGCTGATTGGAACAGGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   229 GCCCATCAGGGCCGGTTGTAATATTCACAATTGGAAGTAAGCTGATTGGAACAGGA 170

Qy   241 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGCAGAGAATCCTTAAATTCAAGA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   169 AAATTATCAGAAGATGAAGAAACTGAAGTACAAGCAGAGAATCCTTAAATTCAAGA 110

Qy   301 TGGGAATGCCTCAGGGTAGCTAGCAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   109 TGGGAATGCCTCAGGGTAGCTAGCAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATG 50

Qy   361 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAA 409
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    49 GATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAA 1
```

```
Qy  1346  TCATGAGTACTCTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAG  1405
                ||||||||||||||||||||||||||||| ||||| |||||||||||||||||||||
Db   187  TCATGAATGCACTTGAGACTGTGCGACTCTTCCTGCGATCAACCAGTAGAGGGACTGG   246

Qy  1406  AGAAACTACCACAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTCACTC  1465
                ||||| ||||||||||  |||| ||||||||||||||||||||||||||||||||||
Db   247  AAAAGGTCTATCCAGAACCAAGAGACCTATCACCTGAGGAGAGGGCCCAGAATGTCACTA   306

Qy  1466  GGCTTCACGAAAGCAGGCTGCGTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACT  1525
                |||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   307  AGTTCTCCGAAGGCAAGCAAGCAGCGATGAGTGTTCAGAACTGAGTGGGATAAGCTAAATCTTACGTT   366

Qy  1526  CCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTGAAAGACTTCAGAGACTTCAAGAGG  1585
                 |||||| ||||||||||||||| ||| |||||||||||||||| |||||||||||
Db   367  CTGCTGATTGGCAAAAGAAGATAGATGTCTTGAAAGACTGCAGGGCTTCAGGAG   426

Qy  1586  CCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGTGAGTCAAGGGATCCTGGCAGC  1645
                |||||||||||||||||||||||||| |||| |||||||||||||||||||||||||
Db   427  CAATGGATGAACTAGACCTGAAACTGCCAGGCTGAAGCATTCAAGGGATCCTGGCAGC   486

Qy  1646  CCGTGGGCGATCTCCTCATTGACTCTCTCCCAAGATCACCTCGAGAGAAGTCAAGGCACTTC  1705
                || |||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   487  CAGTGGGGGATCGTGCTGATAGACTCTCTGCGAGGATCACTTAGAAAAAGTCAAGGTTTATC   546

Qy  1706  GAGGAGAAATTGCGCTCTGAAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGC  1765
                ||||| || ||| |||||||| |||| |||||||||| ||||| || |||||||||
Db   547  GAGCAGAAATGGTGCCCCTTANAGAGAAGGTGCATCAAGTCAATGAGCTGGCTCACCGGT   606

Qy  1766  TTACCACTTGGCATTCAGCTTCCACCGTATAACCTC  1803
                 |||  |||||||  ||||||||| |||||||||
Db   607  TCGCTCCCCCTGATATTCAGTTCTCCCCATACACTCTC   644


RESULT 14
CB547284/c
LOCUS       CB547284    AMGNNUC:SRPB2-00242-E3-A srpb2 (10220) Rattus norvegicus cDNA clone
DEFINITION  AMGNNUC:SRPB2-00242-E3-A srpb2 (10220) Rattus norvegicus cDNA clone
            srpb2-00242-e3 5', mRNA sequence.      402 bp    mRNA      linear    EST 01-APR-2003
ACCESSION   CB547284
VERSION     CB547284.1  GI:29431225
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 402)
AUTHORS     Amgen EST Program.
TITLE       Amgen Rat EST Program
JOURNAL     Unpublished (2003)
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00242  row: e  column: 3.
FEATURES             Location/Qualifiers
     source          1..402
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="srpb2-00242-e3"
                     /tissue_type="prostate tissue"
                     /clone_lib="srpb2 (10220)"
                     /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; rat
                     prostate normalized double selected poly(A+) mRNA size
                     fraction > 1 kb"
ORIGIN

Query Match     16.2%;  Score 324.6;  DB 14;  Length 402;
Best Local Similarity   90.8%;  Pred. No. 1.4e-62;
Matches   345;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy  1236  TGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAA  1295
```

```
RESULT 13
BU313510
LOCUS       BU313510    603540290F1 CSEQCHN61 Gallus gallus cDNA clone CHEST508c24 5', mRNA      644 bp    mRNA      linear    EST 28-NOV-2002
DEFINITION  603540290F1 CSEQCHN61 Gallus gallus cDNA clone CHEST508c24 5', mRNA
            sequence.
ACCESSION   BU313510
VERSION     BU313510.1  GI:25821511
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 644)
AUTHORS     Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE       A Comprehensive Collection of Chicken cDNAs
JOURNAL     Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE     22335534
PUBMED      12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1..644
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Compton line 151"
                     /db_xref="taxon:9031"
                     /clone="CHEST508c24"
                     /sex="Female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN61"
                     /note="Organ: heart; Vector: pBluescript II KS(+); Site_1:
                     EcoRI; Site_2: NotI; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
ORIGIN

Query Match     19.4%;  Score 389;  DB 13;  Length 644;
Best Local Similarity   75.5%;  Pred. No. 4.4e-77;
Matches   482;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy  1166  CCAGTTCTCGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGCAGGAACTTCTGGTGTGGCTAC  1225
                |||  |||||||  |||||||||  |||| ||||| |||||| ||||||| |||||||
Db     7  CGAGCACAGACCAGTGGAAGCCGTTACATCTCTCTTCAGGAACTTTGGCATGCTGC   66

Qy  1226  AGCTGAAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTC  1285
                |||||| || |||||||||||||||||| ||||||||||||||||||||||||||||
Db    67  AATTGAAGGAGGATGAATTAAAACAGCAAGCACCCCATTGGTGAGATATTCCCACTGTGC   126

Qy  1286  AGAAGCAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAGAAACCTGTAA  1345
                ||||||||| || |||| |||||||||||||||||||  ||||| |||||||||||||
Db   127  AGAAGCAGAATGATGTTCATAGGACTTTCAAGAGGAGGAGCTGAAAAGAACAAAAGAACCTGTTA   186
```

```
Qy   1104  CATGAACTTCAAGTGGACTGGAACTTCGGAGAAAAAGTCTCCAACATTAGTTCCCATTTGGA  1163
                ||||||||        |||||   |||                   ||| |  ||||  ||||
Db    234  CATGAAATCTGTGTTGGAATGAAATAAAAAAAGAAGTCTCACAACCTCCGCGCTCGCCTAGA   293

Qy   1164  AGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGCT    1223
                |  ||   |||  |||||||||||  ||||||    |||||||||        |  |||||
Db    294  GGCCTTCTCAGACACCAGTGGAAAGCTTCAGCGCCCCTTCAAGAGATTATTGACTGGCT     353

Qy   1224  ACAGTCGAAAGATGATGAATTAAGCCGGCAGCGGCTATTGGAGGGACTTCCAGCAGT     1283
                ||| |  |||||||||| |||||   |||||| |||   ||||||||||||||||||
Db    354  CAGCCAAAAGGATGAGAGTTGTCAGTCAGTTGTCAGTCAGTCGGCGGGGATGTGGCCCTGGT   413

Qy   1284  TCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGT    1343
                |||||    ||||||   |||       ||||| ||||||||||||          |||
Db    414  GCAACAGGAGAAGGAGCACATGGCGCGAATGCGCGGGGGCCCTA                    473

Qy   1344  AATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCAGCCCTTTGGAAGGACT  1403
                |||||||||| |||||||||||| |||||||||||||||||||||||   ||||||||||||
Db    474  CATCTATTCTGTGTGGACTGGACTCAGACTGGACTGGACTGCAGCCCAGTGGAGTT          533

Qy   1404  AGAGAAACTCTACCAGGAGCCCAGAGAGTCGCTGCCTCCTGAGGAGAGCCCAGAATGTCAC  1463
                ||||||||||||       |||||||||||||  |||||||||||||||||||||||  |
Db    534  AGAGAGCCTCATTCTGAGCAGCAAAGATACCTCCCCGAAACAGCCGGATCCAGAATCTCAG    593

Qy   1464  TCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCA  1523
                |  |||||||||||||||||||| |||||||||||||||||||||  ||||||||||||
Db    594  CCGCTTTGTATGGAAGCAGGCGGCAGGTGGTGCGTGGAGAAGTTGACAGCGCCCG         653

Qy   1524  CTCCCGTGACTGGCGAGAGAAAATAGATGAGACCCCTTGAAAGACTTCCAGGAACTTCAAGA  1583
                |||| ||  ||||| ||||||| ||||||||||||||||||||||||||||||||||||
Db    654  CTGTGTGGACCAGCACCGTCACATTGAGCGGACTCTGGAGCAGCTCTTGGAGATTCAAGG    713

Qy   1584  GGCCACGGATGAGCTGGACCTCAGCTGGAGCGACTCGGGA                        1643
                |||| |||| |||||||||||||||||||||||  |||||||||
Db    714  GGCAATGAGGAACTAAGCGCCTACTACTGAGCTGAGGAGTCGAGGAGTCCAGCCACTTGGGA   773

Qy   1644  GCCCGTGGCGGCGATCTCCTCATGACTCTCTCCAAGATCACCTGAGAAAGTCAAGGCACT    1703
                |||||||||| ||||||||||||||| |||| ||||||||||||| |||||||| |||
Db    774  GCCCATTGGGGATCTCTTCATTGATTCACTCCCAGAGCAGCACATCCAGGCTATTAAGCTGTT   833

Qy   1704  TCGAGGAGAAATTGCGCCTCTGAAAGAGAGAGTGAGCCAGTGCCATTGTCATGACCTTGCTGCCA  1763
                |||||||||| ||||  ||||||||||||||||||||||||||||||||  |||||||||||
Db    834  CAAAGAAGAATTCTCCCCCATGAAAGATGGAGTAAAGTTGGGTAATGATCTGGCCCACCA    893

Qy   1764  GCTTACCACTTGGGCATTCAGCTCTGCACTTGTCAATGGAGAATTCCACGGCCCTGGAAGACCTGAA  1823
                ||||||||||||||||||||||||||||||  |||| |||||||||||||||||||||||
Db    894  ACTTGCCATTCTGATGTGCACTTGTCACTTGTCAATGGGAACCCTGGAACACAGATCAA    953

Qy   1824  CACCAGATGGAAGCTTCTGCAGTGGCCGTGCAGGAACCGAGTCAGGCAGCTCATGAAGC    1883
                ||||||||||||| |||||||  |||||| || |||||||||| ||||||||||||||||
Db    954  CGTCCGATGGAAACAACTACACAGGCGTCAGTTAGTGAGAGCTCCAGATGC          1013

Qy   1884  CCACAGGGACTTTGGTCGTCAGCATCTCAGCACTTTCTTCCACGTCTGTCAGGGTCCCTG  1943
                |||||||||||| |  ||||||||||||||||||||||  |||||||||||||||||||
Db   1014  CCACCGGGGACTTTGGGGTCCACCCTGGGTCACAGCACTTTCTCCTCCTCCTCCTGTCCCCTG  1073

Qy   1944  GGAGAGCAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAA                     1985
                |||| || | |||||| ||||||| ||||||||||| ||||||
Db   1074  GGAAAGAGCAATTTCACCCAATAAAGTTCCCTACTACTACATCAA                   1115
```

Search completed: April 5, 2004, 08:15:04
Job time : 3501.87 secs

```
Db    402  TGATGAGTTGAGCGGCAGGCACCTATTGGTGGCGATTTTCCAGCAGTTCAGAGCAGAA   343

Qy   1296  CGAATGTACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTAC  1355
                ||| |||||| ||||||||||||||||||||||  ||||||||||||||||||||||||||
Db    342  TGATGTACACAGGGCCCTTCAAGAGGGAATTGANAACTAAAGAACCTGTAATCATGAGTAC   283

Qy   1356  TCTTGAGACTGTACGAATATTTCGACAGAGCAGCGCCTTTGGAAGGACTAGAGAAACTCTA  1415
                |||||||||||| ||||||||||  ||||||||||  ||||||||||||||||||||||||
Db    282  TCTTGAGACTGTGAGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAAACTCTA   223

Qy   1416  CCAGGAGCCCAGAGAGTCTGCCTCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCTACG  1475
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db    222  CCAGGAGCCCAGAGAGTCTGCCTCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTCCTACG   163

Qy   1476  AAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTGCGCTGACTG  1535
                ||||||||||||| ||||||| |||||||||||| |||||||||||||||||||| ||| |
Db    162  AAAGCAGGCTGAACGAGGTCAACACTGAATGGGACAAATTGAACCTGCACCTGCAGCTGATTG   103

Qy   1536  GCAGGAAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGGATGA   1595
                |||| ||||||||||||||||||  ||||||||||| ||||||||  ||||  ||||||||
Db    102  GCAGCGAAAAATAGATGAGACCCTCTTGAAAGACTTCAGGAACTTCAGGAAGCTGCCGGATGA    43

Qy   1596  GCTGGACCTCAAGCTGCGCC                                            1615
                ||| |||||||| || |||
Db     42  ACAGCCACCTCATGTTGCGGC                                             23
```

RESULT 15
AY408546

LOCUS   AY408546   2874 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION Homo sapiens DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY408546
VERSION   AY408546.1 GI:39764517
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2874)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 2874)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..2874
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>2874
                     /gene="DRP2"
                     /locus_tag="HCM3260"
ORIGIN

Query Match   15.5%;   Score 310.8;   DB 29;   Length 2874;
Best Local Similarity   59.5%;   Pred. No. 2.8e-59;
Matches 525;   Conservative 0;   Mismatches 357;   Indels 0;   Gaps 0;

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: April 4, 2004, 12:04:07 ; Search time 6259.59 Seconds
                                 (without alignments)
                                 16625.161 Million cell updates/sec

Title: US-09-845-416-9_COPY_600_3000
Perfect score: 2401
Sequence: 1 acatgcattcaacatcgcca........tggatctcctgagcctgtca 2401

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
      10: gb_ro:*
      11: gb_sts:*
      12: gb_sy:*
      13: gb_un:*
      14: gb_vi:*
      15: em_fun:*
      16: em_hum:*
      17: em_in:*
      18: em_mu:*
      19: em_om:*
      20: em_or:*
      21: em_ov:*
      22: em_pat:*
      23: em_ph:*
      24: em_pl:*
      25: em_ro:*
      26: em_sts:*
      27: em_un:*
      28: em_vi:*
      29: em_htg_hum:*
      30: em_htg_inv:*
      31: em_htg_other:*
      32: em_htg_mus:*
      33: em_htg_pln:*
      34: em_htg_rod:*
      35: em_htg_mam:*
      36: em_htg_vrt:*
      37: em_sy:*
      38: em_htgo_hum:*
      39: em_htgo_mus:*
      40: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	1839.4	76.6	5462	6	AX538621	AX538621 Sequence
2	1725	71.8	5339	6	AX538620	AX538620 Sequence
3	1419	59.1	5417	6	AX538619	AX538619 Sequence
4	1335.6	55.6	5952	6	AR304538	AR304538 Sequence
5	1335.6	55.6	5952	6	AX114289	AX114289 Sequence
6	1335.6	55.6	8689	6	AX538622	AX538622 Sequence
7	1335.6	55.6	11443	6	AX538624	AX538624 Sequence
8	1335.6	55.6	12057	6	AX538627	AX538627 Sequence
9	1335.6	55.6	12446	9	HSDMDR	X14298 Human mRNA
10	1335.6	55.6	13957	6	AX409637	AX409637 Sequence
11	1335.6	55.6	13957	6	AX538581	AX538581 Sequence
12	1335.6	55.6	13957	9	HUMDYS	M18533 Homo sapien
13	1335.6	55.6	13977	6	AR220819	AR220819 Sequence
14	1179.4	49.1	13887	4	AF070485	AF070485 Canis fam
15	1143.6	47.6	13815	6	AX306153	AX306153 Sequence
16	1143.6	47.6	13815	6	AX538582	AX538582 Sequence
17	1143.6	47.6	13815	10	MUSDYSA	M68859 Mouse dystr
18	1143.6	47.6	19307	6	AR093392	AR093392 Sequence
19	1143.6	47.6	19307	6	AR142592	AR142592 Sequence
20	1093.8	45.6	4402	6	E30220	E30220 Shortened d
21	1021.2	42.5	4402	6	E30219	E30219 Shortened d
22	1018.8	42.4	4075	6	E30221	E30221 Shortened d
23	879.6	36.6	13575	5	GGDYS	X13369 Chicken mRN
24	814.6	33.9	1481	9	BC009242	BC009242 Homo sapi
25	709.2	29.5	3747	6	E30218	E30218 Shortened d
26	644.4	26.8	3275	10	MUSDYS	M18025 Mouse dystr
27	530.2	22.1	3521	5	AF339031	AF339031 Danio rer
28	498.6	20.8	10705	10	RNAJ02967	AJ002967 Rattus no
29	495.6	20.6	11096	6	AX538584	AX538584 Sequence
30	495.6	20.6	11096	10	MMY12229	Y12229 M.musculus
31	492.2	20.5	2654	10	FSCDYSTRO	M37645 Torpedo cal
32	487	20.3	6045	6	A63605	A63605 Sequence 7
33	487	20.3	6045	6	AR281528	AR281528 Sequence
34	487	20.3	6059	6	AX107972	AX107972 Sequence
35	487	20.3	10302	6	AX538583	AX538583 Sequence
36	487	20.3	10302	9	HSMUPS	X69086 H.sapiens m
37	487	20.3	10320	6	A63607	A63607 Sequence 9
38	487	20.3	10320	6	AR281529	AR281529 Sequence
39	476.4	19.8	630	9	HSDMDF1	X06179 Human fetal
40	459	19.1	3161	10	MMGUTRPH	X83506 M.musculus
41	411.8	17.2	3499	9	HSU43519	U43519 Human dystr
42	411.8	17.2	5106	6	AX552248	AX552248 Sequence
43	403.8	16.8	3073	10	AF195788	AF195788 Rattus no
44	403.8	16.8	3140	10	AF195787	AF195787 Rattus no
45	395.8	16.5	4416	10	BC060605	BC060605 Mus muscu

ALIGNMENTS

RESULT 1
AX538621
LOCUS   AX538621   5462 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 41 from Patent WO0229056.
ACCESSION   AX538621
VERSION   AX538621.1   GI:25271168
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS   Chamberlain,J.S. and Harper,S.Q.
TITLE   Mini-dystrophin nucleic acid and peptide sequences
JOURNAL   Patent: WO 0229056-A 41 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

Pred. No. is the number of results predicted by chance to have a

```
FEATURES             Location/Qualifiers
     source          1..5462
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"

ORIGIN

Query Match     76.6%;  Score 1839.4;  DB 6;  Length 5462;
Best Local Similarity  88.1%;  Pred. No. 0;
Matches 2115;  Conservative 0;  Mismatches 76;  Indels 210;  Gaps 2;

Qy      1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 858

Qy     61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCCTTCCA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCCTTCCA 918

Qy    121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAGTGGAAATGTTGCCAAGGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAGTGGAAATGTTGCCAAGGCC 978

Qy    181 ACCTAAAGTGACTAAAGAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 ACCTAAAGTGACTAAAGAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA 1038

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGAGGACAGTTCTTCGCCCCTAAGCCTGATTCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 GATCACGGTCAGTCTAGCACAGGGATATGAGGACAGTTCTTCGCCCCTAAGCCTGATTCAA 1098

Qy    301 GAGCTATGCCTACACACAGGCGTGCTTATGTCACCACCTCACGGAGCCCATT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCGTGCTTATGTCACCACCTCACGGAGCCCATT 1158

Qy    361 TCCTTCACAGATTTGGAAGCTCCTGAAGACAGTCATTGGCAGTTCATTGATGGAGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGATTTGGAAGCTCCTGAAGACAGTCATTGGCAGTTCATTGATGGAGAG 1218

Qy    421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1278

Qy    481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTCGGAGTGGTGAAAGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTCGGAGTGGTGAAAGA 1338

Qy    541 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1398

Qy    601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy    661 AACTGAAGTACAAGAGCAGAGATGAAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGAGATGAAATCTCCTAAATTCAAGATGGGAATGCCTCAGGTAGC 1518

Qy    721 TAGCAGAGTTGAATGCAAGCAAGCAATTACATAGAGTTTAATGGATCTTCCAGAATCAGAAACT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 TAGCAGAGTTGAATGCAAGCAAGCAATTACATAGAGCTCCTGGACTG 1558

Qy    781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAAGAACAAGGAAAATGGAGGAAGA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1559 --------------------------------------------------AC 1560

Qy    841 GCCTCTTGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACAGCCAGTACAACACATAAGGTGCTTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CACTATTGGAGCCCTCTACTCAGACTGTTACTCTGGTGACACAACCTGTGTTACTAA 1620

Qy    901 AGAAGATCTAGAACAAGAACAGTCAGGGTCAATTCTCTCACTCAACATGGTGTGTAGT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ... 

Db   1621 GGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATGTTGGAG---------- 1669
          |||
Qy    961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAGAGAACAACTTAAGGTATTGG 1020

Db   1670 ------------------------------------------------------------ 1669
Qy   1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAGACCGCTGGGTTCTTTTACAAGACAC 1080

Db   1670 ------------------------------------------------------------ 1669
Qy   1081 TCATGATTACTGCAACAGTTCCCCTGCAACCTGGAAAAGTTTCTTGCTGGCTTACAGA 1140

Db   1670 -CATAGATTACTGCAACAGTTCCCCTGCAACCTGGAAAAGTTTCTTGCTGGCTTACAGA 1728
Qy   1141 AGCTGAAACAACTGCCAATGTCCTCAGGATGTCCTAAGGATGGCTACCGTAAGGAAAAGGCTCCTAGAGAGA 1200

Db   1729 AGCTGAAACAACTGCCAATGTCCTCAGGATGTCCTAAGGATGGCTACCGTAAGGAAAAGGCTCCTAGAGAGA 1788
Qy   1201 CTTCCAAGGGAGTAAAAGAGCTGATGAAACAAATGGCAAGCAAGACCTCCAAGGTGAAATTGAAGC 1260

Db   1789 CTTCCAAGGGAGTAAAAGAGCTGATGAAACAAATGGCAAGCAAGACCTCCAAGGTGAAATTGAAGC 1848
Qy   1261 TCACACAGATGTTATCACAACCTGATGATGAAAACAGCCAAAAAATCCTGAGATCCCTGA 1320

Db   1849 TCACACAGATGTTATCATCAACCTGATGATGAAAACAGCCAAAAAATCCTGAGATCCCTGA 1908
Qy   1321 AGGTTCCGATGATGAGTCCTGTTACAACAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380

Db   1909 AGGTTCCGATGATGCGTCCTGTTACAACAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1968
Qy   1381 TGAACTTCGGAAAAAGTCTCTCAACATTTGGAAGCCAGTGGTCTCCATTTGGAAGGAGCCCATT 1440

Db   1969 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTGGTCTCCATT 2028
Qy   1441 GAAGCGTCTGCACCTTTCTCTGCGAGGAACTTCTCGGTGTGGTGCTACAGCTCAAGATGATGA 1500

Db   2029 GAAGCGTCTGCACCTTTCTCTGCGAGGAACTTCTCGGTGTGGTGCTACAGCTCAAGATGATGA 2088
Qy   1501 ATTAAGCCGGCAGCACCTATTGGAGGAGCCTTCGACTTCAGCAGTTCGAAGCAGAACGATGT 1560

Db   2089 ATTAAGCCGGCAGCACCTATTGGAGGAGCCTTCGACTTCAGCAGTTCGAAGCAGAACGATGT 2148
Qy   1561 ACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620

Db   2149 ACATAGGGCCTTCAAGAGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2208
Qy   1621 GACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAGGACTAGAGAAACTCTACCAGGA 1680

Db   2209 GACTGTACGAATATTTCTGACAGAGCAGCCCTTGGAAGGACTAGAGAAACTCTACCAGGA 2268
Qy   1681 GCCCAGAGAGCTGCCTCCTCCTGGAGGAGGAGCCCAGAATGTCACTCGGCCTTCACTCGGAAAAGCA 1740

Db   2269 GCCCAGAGAGCTGCCTCCTCCTGGAGGAGGAGCCCAGAATGTCACTCGGCCTTCACTCGGAAAAGCA 2328
Qy   1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGACACTGCACTCCGCTGACTGGCAGAG 1800

Db   2329 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGACACTGCACTCCGCTGACTGGCAGAG 2388
Qy   1801 AAAAATAGATGAGACCCTTGAAAGACTTCAAGGAACTTCAAGAGGGCCACCGATGAGCTGGA 1860

Db   2389 AAAAATAGATGAGACCCTTGAAAGACTTCAAGGAACTTCAAGAGGGCCACCGATGAGCTGGA 2448
Qy   1861 CCTCAAGCTGCGCCAAGCTGAGTGTCAAGGATCTGGCGATCCTGACTGGCAGATCTCCT 1920

Db   2449 CCTCAAGCTGCGCCAAGCTGAGTGTCAAGGATCTGGCGATCCTGACTGGCAGATCTCCT 2508
Qy   1921 CATTGACTTGCTCTCCAAGATCACCTCCAAGGCACTTCGAGGACGAAAGTTGCGCC 1980

Db   2509 CATTGACTTGCTCTCCAAGATCACCTCCAAGGCACTTCGAGGAGAAAGTTGCGCC 2568
Qy   1981 TCTGAAAGAACGTGAAGCCACGTCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGCAT 2040
```

```
Db   2569 TCTGAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTCTTACCACTTTGGGCAT 2628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2629 TCAGTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2688

Qy   2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGGCCACAGGACTTTGGTCC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2689 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGGCCCACAGGACTTTGGTCC 2748

Qy   2161 AGCATCTCAGCACTTCTTCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2749 AGCATCTCAGCACTTCTTCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2808

Qy   2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTGCTGGGACCATCCCAA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2809 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTGCTGGGACCATCCCAA 2868

Qy   2281 AATGACAGAGCTCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2869 AATGACAGAGCTCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2928

Qy   2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2929 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2988

Qy   2401 A 2401
          |
Db   2989 A 2989
```

```
Db    241 GATCACGGTCAGTCTAGCACACAGGGATATGAGAGAATGAGAGAGAACTTCTTCCCCTAAGCCTCGGATTCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1039 GATCACGGTCAGTCTAGCACACAGGGATATGAGAGAATGAGAGAGAACTTCTTCCCCTAAGCCTCGGATTCAA 1098

Db    301 GAGCTATGCTACACACAGGGCTGCTTATGTCCACCACCTCTGACCTCTCACCAGATGGAAGCCCATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1099 GAGCTATGCTACACACAGGGCTGCTTATGTCCACCACCTCTGACCTCTCACCAGATGGAAGCCCATT 1158

Db    361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

Db    421 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATTATCGTGCTTCTTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1219 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATTATCGTGCTTCTTTC 1278

Db    481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA 1338

Db    541 CCAGTTTCATACTCATGAGGGTACATGATTTGACACGCCCATCAGGGACCGGGTTGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1339 CCAGTTTCATACTCATGAGGGTACATGATTTGACAGCCCATCAGGGGCCGGGTTGG 1398

Db    601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGGAACAGGAGAAAATTATCAGAAGATGAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGGAACAGGAGAAAATTATCAGAAGATGAAGA 1458

Db    661 AACTGAAGTACAAGACCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGGAATGCCTCAGGGTAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1459 AACTGAAGTACAAGACCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGGAATGCCTCAGGGTAGC 1518

Db    721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 780
          ||||||||||||||||||||||||||||||-----------------
Qy   1519 TAGCATGGAAAAACAAAGCAATTTACA------------------------ 1545

Db    781 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAGAGAAAGAACAAGGAAAAATGAGGAAGAAGA 840
          ----------------------------------------------------
Qy   1546 ---------------------------------------------------------- 1545

Db    841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACAATAAGGTGCTTCA 900
          ------------------------------------------------------------
Qy   1546 ---------------------------------------------------------- 1545

Db   1545 --------------------------------------------------------- 1545
Qy    901 AGAAGGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTCGTGGTGGTAGT 960
          ------------------------------------------------------------
Db   1546 ---------------------------------------------------------- 1545

Qy    961 TGATGAATCTAGTGGAGATCACCGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
          ------------------------------------------------------------
Db   1546 ---------------------------------------------------------- 1545

Qy   1021 AGATCGATGGGCAAACATCTGTAGATGACGAAGAACCGCTGGGTTCTTTTACAAGACACAC 1080
          ------------------------------------------------------------
Db   1546 ---------------------------------------------------------- 1545

Db   1081 TCATAGATTACTGCAACAGTTCCCCTGCAACAGTTCTTGCTCGGCTTACAGA 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1546 TCATAGATTACTGCAACAGTTCCCCTGCAACAGTTCTTGCTCGGCTTACAGA 1605

Qy   1141 AGCTGAAACAACTGCCAAATGCTCCATCTTAATGTACATCACATCACTCCTAGAAGA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 AGCTGAAACAACTGCCAAATGCTCCATCTTAATGTACATCACATCACTCCTAGAAGA 1665

Qy   1201 CTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGAAATTGAAGC 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 CTCCAAGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGAAATTGAAGC 1725

Qy   1261 TCACACAGATGTTTATCACAACCTGATGAAAAATCTGATAAAAATCCTGAGATCCCTGA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1726 TCACACAGATGTTTATCACAACCTGATGAAAAATCTGATAAAAATCCTGAGATCCCTGA 1785
```

```
RESULT 2
AX538620
LOCUS       AX538620                5339 bp    DNA        linear   PAT 23-NOV-2002
DEFINITION  Sequence 40 from Patent WO0229056.
ACCESSION   AX538620
VERSION     AX538620.1  GI:25271166
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 40 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..5339
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3630"
                     /note="Synthetic"
ORIGIN
```

```
Query Match        71.8%;  Score 1725;  DB 6;  Length 5339;
Best Local Similarity  86.1%;  Pred. No. 0;
Matches 2068;  Conservative  0;  Mismatches  0;  Indels 333;  Gaps  1;

Qy      1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 858

Qy     61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 918

Qy    121 AGTTTTGCTCAACAAGTGCATTGAAGCATTGAGCACCAAGGAAGTCGAAAGTGTTGCCAAGGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 AGTTTTGCTCAACAAGTGCATTGAAGCATTGAGCACCAAGGAAGTCGAAAGTGTTGCCAAGGCC 978

Qy    181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAAATGCACTAATTCTCAACA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAAATGCACTAATTCTCAACA 1038
```

```
Db                    2866 A 2866
                            |

RESULT 3
AX538619
LOCUS       AX538619           5417 bp    DNA        linear   PAT 23-NOV-2002
DEFINITION  Sequence 39 from Patent WO0229056.
ACCESSION   AX538619
VERSION     AX538619.1  GI:25271163
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 39 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..5417
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

Query Match             59.1%;  Score 1419;  DB 6;  Length 5417;
Best Local Similarity   78.1%;  Pred. No. 0;
Matches 1876;  Conservative   0;  Mismatches  270;  Indels  255;  Gaps    4;
```

```
Qy     1 ACATGCATTCAACATCGCCAGATATCAATTAGCGCATAGAGAAACTACTCGATCCTGAAGA 60
         ||||||||||||||||||||||||||  ||| |  ||||||||||||||||||| |||||||
Db   799 ACATGCATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAGAAACTACTCGATCCTGATCCTGAAGA 858

Qy    61 TGTTGATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
         ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db   859 TGTTGATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 918

Qy   121 AGTTTGCCTCACAAGTGAGCATTGAGCCATCCAGGAGAATGTGCCAAGGCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   919 AGTTTGCCTCAACAAGTGAGCATTGAGCCATCCAGGAGAATGTGCCAAGGCC 978

Qy   181 ACCTAAAGTGACTAAAGAAAGAAGACAATTTCAGTTACATCAATCATCAAATGCACTAATTCTCAACA 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   979 ACCTAAAGTGACTAAAGAAAGAAGACAATTTCAGTTACATCAATCATCAAATGCACTAATTCTCAACA 1038

Qy   241 GATCACGGTCAGTCTAGCACAGGGATATGAGAATGAGACACAGGGATATGAGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1039 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1098

Qy   301 GAGCTATGCTACACAGAGGCTGCTTATGTCACCACCTCACCACCTCATTGGCAGTCATTGAGGAGCCCATT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 GAGCTATGCTACACAGAGGCTGCTTATGTGACCACCTCACCACCTCATTGGCAGTCATTGATGAGGAGCCCATT 1158

Qy   361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGAGAGAAGCATTCATTGGCAGTTCATTGATGGAGAG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGAGAGAAGCATTCATTGGCAGTTCATTGATGGAGAG 1218

Qy   421 TGAAGTAAACCTGACCGTTATCACAAACAGCTTGCTGGGACCATCCCAA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 TGAAGTAAACCTGACCGTTATCAAACAGCTTGCTGGGACCATCCCAA 1278

Qy   481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGTGAAAGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGTGAAAGA 1338

Qy   541 CCAGTTTCATCATGGGTACATGAGAGGGTACATGAGGAGTGGTCCGGGTTGG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTTCATCATGGGTACATGAGGGGTACATGAGGAGTGGTCCGGGTTGG 1398

Qy   601 TAATATTCCACAATTGGGAAGTAAGCTGATTGGAACAAGAAAATTATCAGAAGATGAAGA 660
```

... (continuation of alignment blocks Qy/Db) ...

```
Qy  1321 AGGTTCCGATGATGCAGTCCTCGTTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380
Qy  1786 AGGTTCCGATGATGCAGTCCTCGTTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1845
Qy  1381 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1440
Db  1846 TGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1905
Qy  1441 GAAGCGTCTGCACCTTCTGCTCGAGGAACTTCTGGTTGTGGCTACAGCTGAAAGATGATGA 1500
Db  1906 GAAGCGTCTGCACCTTCTGCTCGAGGAACTTCTGGTTGTGGCTACAGCTGAAAGATGATGA 1965
Qy  1501 ATTAAGCCGGCAGGCACCTATTGGAGCGACTTCCAGCAGTTCAGAAGCAGAACGATGT 1560
Db  1966 ATTAAGCCGGCAGGCACCTATTGGAGGGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 2025
Qy  1561 ACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
Db  2026 ACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2085
Qy  1621 GACTGTACGAACATATTTCTGACAGAGCAGCTAGAGAACTTCACCAGGA 1680
Db  2086 GACTGTACGAACATATTTCTGACAGAGCAGCTAGAGAAACTCCTACCAGGA 2145
Qy  1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGCGCCCAGAATGTCACTCGGCTTCACGGAAAGCA 1740
Db  2146 GCCCAGAGAGCTGCCTCCTGAGGAGAGCGCCCAGAATGTCACTCGGCTTCACGGAAAGCA 2205
Qy  1741 GGCTGAGGAGGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
Db  2206 GGCTGAGGAGGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2265
Qy  1801 AAAAATAGATGAGACCCTTGAAAGACTTCAAGGCCACGGATGAGCTGGA 1860
Db  2266 AAAAATAGATGAGACCCTTGAAAGACTTCAAGGCCACGGATGAGCTGGA 2325
Qy  1861 CCTCAAGTGTGCGCCAAGCTGAGTGAGTCACTCCAGGGCGCCCGTTGCGCGCAGGGCGATCTCCT 1920
Db  2326 CCTCAAGTGTGCGCCAAGCTGAGTGAGTCACTCCAGGGCGCCCGTTGCGAGGCGCGATCTCCT 2385
Qy  1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCACTTCGAGGAGAATTGCGCC 1980
Db  2386 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCACTTCGAGGAGAAATTGCGCC 2445
Qy  1981 TCTGAAAGAGAAGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2040
Db  2446 TCTGAAAGAGAAGTGAGCCACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2505
Qy  2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2100
Db  2506 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2565
Qy  2101 GCGAGGTGGCCGTCGAGGACCGAGTCAGGCAGTCAGGCGCATGAAAGCCCCAGAGGGGACTTTGGTCC 2160
Db  2566 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTCAGGCGCATGAAAGCCCCAGAGGGGACTTTGGTCC 2625
Qy  2161 AGCATCTCAGCACTTCTTTCCACGTCTGTCCACGTCTTCTTTCCACGTCTGTCCTCGGGAGGAGAGCCATCTCGCC 2220
Db  2626 AGCATCTCAGCACTTCTTTCCACGTCTGTCCTCGGGAGGAGAGCCATCTCGCC 2685
Qy  2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA 2280
Db  2686 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA 2745
Qy  2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
Db  2746 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2805
Qy  2341 GACTGCCATGAAACTCCGAAGACTCAGAAGACTCAGAGGGACCTTTGCTTGAAGTCTTTGCTTGAAGTCTTTGAGCCTGTC 2400
Db  2806 GACTGCCATGAAACTCCGAAGAGACTCAGAGGGACCCTTTGCTTGAAGTCTTTGAGCCTGTC 2865
Qy  2401 A 2401
```

```
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGAGAAATTATCAGAGATGAAGA 1458
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAAGTGCCTCAGGGTAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy    721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy    781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAGGAGAAAATGGAGGAAGA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAGGAGAAAATGGAGGAAGA 1638

Qy    841 GCCTCTTGGACCTGATCTTGAAGACTTAACGGCCAAGTACAACACAACATAAGGTGCTTCA 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACTTAACGGCCAAGTACAACACAACATAAGGTGCTTCA 1698

Qy    901 AGAAGATCTAGAACAAGAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGGTAGT 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAGCAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGGTAGT 1758

Qy    961 TGATGAATTCTAGTGGAGATCACGCAACTTGCTGCTTGGAAGACAACTTAAGGTATTGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATTCTAGTGGAGATCACGCAACTTGCTGCTTGGAAGACAACTTAAGGTATTGGG 1818

Qy   1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGACAC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   1081 TCATAGATTACTGCAACAGTTCCCCTGACTGGAAAAGTTCTTGCTGGCTTACAGA 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGCACAGTCTTACTGAAGACAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   1141 AGCTGAAAACAACTGCCAATGTCCTACAGAGATGTCCTACACCGGTAAGGAAAGGCTCCTAGAAGA 1200
           ||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAGAAGATTCACCACAACTGGCTTTA--------AAGA 1983

Qy   1201 CTCCAAGGAGTAAAAGAGTGATGAAACAATGGCAAGACCTCAAGTGAAATTGAAGC 1260
           |||||||||||||||||||||||||||||||||||||
Db   1984 TCAAAATGAAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAA 2043

Qy   1261 TCAACAGATGTTTATCACAAACTGGATGAAAAACAGCCAAAAAAATCCTGAGATCCCTGGA 1320
           |||||||||||||||||||||||||||
Db   2044 GAAAAAGCAATCCATGGGCAAACGTATTCCATCCAAACAAGATCTTCTTTCAACACTGAA 2103

Qy   1321 AGGTTCCGATGATGCAGTCCTGTTACAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380
           ||||||||||||||||
Db   2104 GAATAAGTCAGTGACCCAGAAGACGGAAGCATGGCTGGGATAACTTTGCCCGGTGTTGGGA 2163

Qy   1381 TGAACTTCGGGAAAAGTCTCTCCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTG 1440
           ||||||||||||||
Db   2164 TAATTTAGTCCAAAAACTTGAAAAGAGTACAGCACA---------------------- 2199

Qy   1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCGTGTGTGGCTACAGTGAAAGATGATGA 1500
           |||||||
Db   2200 --------------------------------------------------------- 2199

Qy   1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 1560
           ||||||
Db   2200 --------------------GATTTCACAGGCTGTCACCACCACTCAGCCAT 2231

Qy   1561 ACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCCTTGA 1620
           |||||
Db   2232 CACTAACACAGCACAACTGTAATGGAAAACAGTAACTACGGTGACCACAGGGGACA----- 2286

Qy   1621 GACTGTTACGGAATATTTCTGACAGAGCAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1680
           ||
Db   2287 --------------------------------------------------------- 2286

Qy   1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAATGTCACTCGGCTTCTACGAAAGCA 1740
           ||||||
Db   2287 ---------------------------GATCCCTGGTAAAGCA 2301
```

```
Qy   1741 GGCTGAGGAGGTCAATATACTGAGTGGGAAAAAATTGAACCTGCACTCCGCTGCACTGGCAGAG 1800
           |||  ||||||
Db   2302 TGCTCCAAGAGG-------------AACTTCCACCACCTCCCAAAAGAGAG 2343

Qy   1801 AAAAATAGATGAGACCCTGAAAGATCTCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
           |||||||||||||||||||||||||||||||||||||||||
Db   2344 GCAGATTACTGTGATCTTGGATCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2403

Qy   1861 CCTCAAGCTCGCGCCAAGCTGAGGTGATCAAGGGGATCCTGGCAGCCGGTGGGGCGATCTCCT 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2404 CCTCAAGCTCGCGCCAAGCTGAGGTGATCAAGGGGATCCTGGCAGCCGGTGGGGCGATCTCCT 2463

Qy   1921 CATTGACTTCTCTCCAAGATCACCTCGAGAAGTCAGGCACTTCGAGCAGAAATTGCGCC 1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2464 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAGGCACTTCGAGGAGAAATTGCGCC 2523

Qy   1981 TCTGAAAGAGAAGATGAGCCAACGTCAATGCACTCTGCGCCAGCTTACCCACTTTGGGCAT 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2524 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTCTGCGCCAGCTTACCCACTTTGGGCAT 2583

Qy   2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGAAGACTCTGGAAGATGGAAGCTTCT 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2584 TCAGCTCTCACCGTATAACCTCAGCACTCTGAAGACTCTGGAAGATGGAAGCTTCT 2643

Qy   2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGGGGTCCCACGAGGACTTTGGTCC 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2644 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGGGGTCCCACGAGGACTTTGGTCC 2703

Qy   2161 AGCATCCAGCACTTCTTTCCACGTCTGTCCAGGGTCCTCTGGAGAGAGCCATCTCGCC 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2704 AGCATCCAGCACTTCTTTCCACGTCTGTCCAGGGTCCTCTGGAGAGAGCCATCTCGCC 2763

Qy   2221 AAACAAAGTGCCCTACTATATCACCACGAGACTCAAACAACAACTTGCTGGGACCATCCCAA 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2764 AAACAAAGTGCCCTACTATATCACCACGAGACTCAAACAACAACTTGCTGGGACCATCCCAA 2823

Qy   2281 AATGACAGAGCTCACCAGTCTTTAGCTGACCTGGATTCAGATTCTCAGCTTATAG 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2824 AATGACAGAGCTCACCAGTCTTTAGCTGACCTGGATTCAGATTCTCAGCTTATAG 2883

Qy   2341 GACTGCCATGAAAACTCCGAAGACTGCAGAGAGCCCTTGCTTGGATCTCTTGAGCCGTC 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2884 GACTGCCATGAAAACTCCGAAGACTGCAGAGAGCCCTTGCTTGGATCTCTTGAGCCGTC 2943

Qy   2401 A 2401
           |
Db   2944 A 2944
```

```
RESULT 4
AR304538
LOCUS       AR304538                5952 bp    DNA        linear   PAT 12-JUN-2003
DEFINITION  Sequence 1 from patent US 6544786.
ACCESSION   AR304538
VERSION     AR304538.1  GI:31693691
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 5952)
  AUTHORS   Xiao,X. and Liu,P.X.
  TITLE     Method and vector for producing and transferring trans-spliced
            peptides
  JOURNAL   Patent: US 6544786-A 1 08-APR-2003;
FEATURES             Location/Qualifiers
     source          1..5952
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             55.6%;  Score 1335.6;  DB 6;  Length 5952;
Best Local Similarity   98.3%;  Pred. No. 0;
```

Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
Qy    1028 TGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACACTCATAGA 1087
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2902 TGGAGAAGCATTCATAAAAGGGTGAGTGAGCGAGAGGCCATCTCGGGAAAACTCATAGA 2961

Qy    1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGCTGAA 1147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2962 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGCTGAA 3021

Qy    1148 ACAACTGCCAATGTCCTACAGGATGCCTAAGGAAAAGGCTCCTAGAAGACTCCAAG 1207
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3022 ACAACTGCCAATGTCCTACAGGATGCCTAAGGAAAAGGCTCCTAGAAGACTCCAAG 3081

Qy    1208 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 1267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3082 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 3141

Qy    1268 GATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCC 1327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3142 GATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCC 3201

Qy    1328 GATGATGCAGTCCTGTTACAAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT 1387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3202 GATGATGCAGTCCTGTTACAAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT 3261

Qy    1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCCAGTGGAAGCGT 1447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3262 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCCAGTGGAAGCGT 3321

Qy    1448 CTGCACCTTTCTCTGCAGGAACTTCGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1507
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3322 CTGCACCTTTCTCTGCAGGAACTTCGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 3381

Qy    1508 CGGCAGGCACCTATGGAGGCGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGG 1567
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3382 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGG 3441

Qy    1568 GCCTTCAAGAGGGAATTGAAAACTAAAGACCTGTAATCATGAGTACTCTTGAGACTGTA 1627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3442 GCCTTCAAGAGGGAATTGAAAACTAAAGACCTGTAATCATGAGTACTCTTGAGACTGTA 3501

Qy    1628 CGAATATTTCTGACAGAGCAGCTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGA 1687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3502 CGAATATTTCTGACAGAGCAGCTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGA 3561

Qy    1688 GAGCTGCCTGCCTCCTGAGAGGAGAGCCCAGAATGTCACTCGGCTTCACGAAAAGCAGGCTGAG 1747
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3562 GAGCTGCCTGCCTCCTGAGAGGAGAGCCCAGAATGTCACTCGGCTTCACGAAAAGCAGGCTGAG 3621

Qy    1748 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAATA 1807
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3622 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAATA 3681

Qy    1808 GATGAGAACATCTGTAGATGGACAGAAGACTCCAGGAACTTCAAGAGCCACGGATGAGCTGGAACCTCAAG 1867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3682 GATGAGAACATCTGTAGATGGACAGAAGACTCCAGGAACTTCAAGAGCCACGGATGAGCTGGAACCTCAAG 3741

Qy    1868 CTGCGCCAAGTCGAGGTGATCAAGGATCCGAGCCCGTGGGCGATCCTCATTGAC 1927
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3742 CTGCGCCAAGTCGAGGTGATCAAGGATCCGAGCCCGTGGGCGATCCTCATTGAC 3801

Qy    1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTCTGAAA 1987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3802 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTCTGAAA 3861

Qy    1988 GAGAACGTGAGCCACGTCAATGACTTCGTCGCCAGCTTACCACTTACCACTTGGGCATTCAGCTC 2047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3862 GAGAACGTGAGCCACGTCAATGACTTCGTCGCCAGCTTACCACTTACCACTTGGGCATTCAGCTC 3921

Qy    2048 TCACCGTATAACCTCAGCACTCTCGGAGACTCGAAGCTTCTGCAGGTG 2107
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    3922 TCACCGTATAACCTCAGCACTCTCGGAGACTCGAAGCTTCTGCAGGTG 3981
```

```
Qy    2108 GCCGTCGAGGACCAGTCAGGAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCT 2167
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3982 GCCGTCGAGGACCAGTCAGGAGCTGCATGAAGCCACAGGGACTTTGGTCCAGCATCT 4041

Qy    2168 CAGACTTTCTTCCACGTCGTCCACGTCCTGGGAGAGCCCTCGGAGAGAGCCATCTCGCCAAACAAA 2227
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4042 CAGACTTTCTTCCACGTCGTCCACGTCCTGGGAGAGCCCTCGGAGAGAGCCATCTCGCCCAAACAAA 4101

Qy    2228 GTGCCCTACTATATCAACCACGAGACTCAAACACTTGCTGGGACCATCCCAAAATGACA 2287
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4102 GTGCCCTACTATATCAACCACGAGACTCAAACACTTGCTGGGACCATCCCAAAATGACA 4161

Qy    2288 GAGGTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 2347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4162 GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 4221

Qy    2348 ATGAAACTCCGAGACTGCAGAGGCCCTTGCTTGGATCCTCGTTGGGATCTCTTGAGCCTGTCA 2401
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4222 ATGAAACTCCGAGACTGCAGAGGCCCTTGCTTGGATCCTCTTGGGATCTCTTGAGCCTGTCA 4275
```

RESULT 5

AX114289			
LOCUS	AX114289	5952 bp	DNA linear PAT 11-MAY-2001
DEFINITION	Sequence 1 from Patent WO0129243.		
ACCESSION	AX114289		
VERSION	AX114289.1	GI:14031259	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1
AUTHORS   Paul,X.L. and Xiao,X.
TITLE   Method and vector for producing and transferring trans -spliced
            peptides
JOURNAL   Patent: WO 0129243-A 1 26-APR-2001;
            DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
            SYSTEM OF HIGHER EDUCATION (US)
FEATURES   Location/Qualifiers
     source   1..5952
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature   2897..2898
                     /note="S4 junction site"
     misc_feature   3198..3199
                     /note="S2 junction site"
ORIGIN

Query Match   55.6%; Score 1335.6; DB 6; Length 5952;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
Qy    1028 TGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACACTCATAGA 1087
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2902 TGGAGAAGCATTCATAAAAGGGTGAGTGAGCGAGAGGCCATCTCGGGAAAACTCATAGA 2961

Qy    1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGCTGAA 1147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2962 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGCTGAA 3021

Qy    1148 ACAACTGCCAATGTCCTACAGGATGCCTAAGGAAAAGGCTCCTAGAAGACTCCAAG 1207
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3022 ACAACTGCCAATGTCCTACAGGATGCCTAAGGAAAAGGCTCCTAGAAGACTCCAAG 3081

Qy    1208 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGGTGAAATTGAAGCTCACACCA 1267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3082 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGGTGAAATTGAAGCTCACACA 3141

Qy    1268 GATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCC 1327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3142 GATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCCTGAGATCCCTGGAAGGTTCC 3201
```

RESULT 6

LOCUS   AX538622   8689 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION   Sequence 42 from Patent WO0229056.
ACCESSION   AX538622
VERSION   AX538622.1 GI:25271171
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE   Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 42 11-APR-2002;
             THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
     source   1..8689
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"

ORIGIN

Query Match   55.6%; Score 1335.6; DB 6; Length 8689;
Best Local Similarity   98.3%; Pred. No. 0;
Matches 1350; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy   1028 TGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTACAAGACACTCATAGA 1087
           |||||| |||||||||| |||||| || ||||||| || | ||||| ||||||||
Db   2948 TGGAGAAGCATTCATAAAAGGGTGAGTGAGGGCTGCTTTGGAGAAACTCATAGA 3007

Qy   1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGTCGAA 1147
           |||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||
Db   3008 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGTCGAA 3067

Qy   1148 ACAACTGCCAATGTCCTACGAGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAG 1207
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3068 ACAACTGCCAATGTCCTACGAGATGCTACCCGTAAGGAAAGGCTCCTAGAGACTCCAAG 3127

Qy   1208 GGAGTAAAAGAGTGATGAAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACA 1267
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3128 GGAGTAAAAGAGTGATGAAAACAATGGCAAGACCTCCAAGTGAAATTGAAGCTCACACA 3187

Qy   1268 GATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAGATCCCTGAAGGTTCC 1327
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3188 GATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAGATCCCTGAAGGTTCC 3247

Qy   1328 GATGATGCAGTCCGTTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT 1387
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3248 GATGATGCAGTCCGTTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT 3307

Qy   1388 CGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT 1447
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3308 CGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT 3367

Qy   1448 CTGCACCTTCTCTGCAGGAACTTCTCGTGTGTCGGCACAGTTCAGAGTGAAAGATGATGAATTAAGC 1507
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3368 CTGCACCTTCTCTGCAGGAACTTCTCGTGTGTCGGCTACAGTTCAGAGTGAAAGATGATGAATTAAGC 3427

Qy   1508 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGG 1567
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3428 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGG 3487

Qy   1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 1627
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3488 GCCTTCAAGAGGGAATTGAAAACTAAAGAACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 3547

Qy   1628 CGAATATTTCTGACAGAGACGCCTTTGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGA 1687
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3548 CGAATATTTCTGACAGAGACGCCTTTGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGA 3607

Best Local Similarity 98.3%;　Pred. No. 0;
Matches 1350;　Conservative　0;　Mismatches　24;　Indels　0;　Gaps　0;

```
Qy   1688 GAGGTGCCTCCTGAGGAGAGAGAGCCCAGAATGTCACTCGGCTTCTTACGAAAGCAGGCTGAG 1747
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3608 GAGCTGCCTCCTGAGGAGAGAGAGCCCAGAATGTCACTCGGCTTCTTACGAAAGCAGGCTGAG 3667

Qy   1028 TGGGCAAACATCTGTAGATGGACAAGAGACCGCTGGGTTCTTTACAGACACTCATAGA 1087
          |||||  ||||||||  ||| ||  |||||||||||||||||||||||||||||||||
Db   5702 TGGAGACGCATTCATAAAAGGGTGAGTGAGCGAGGAGGCTGCTTGGAGAAACTCATAGA 5761

Qy   1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAA 1147
          |||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db   5762 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAA 5821

Qy   1148 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAAGGCTCCTAGAAGACTCCCAAG 1207
          |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||||
Db   5822 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAAGGCTCCTAGAAGACTCCCAAG 5881

Qy   1208 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 1267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5882 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 5941

Qy   1268 GATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAGGTGAACTT 1327
          ||||||||||||||||||| ||||||  |||||||||||||||||||||||||||||||||
Db   5942 GATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGAGGTGAACTT 6001

Qy   1328 GATGATGCAGTCCTGTTACAAAGACGTTTGATATACAGCCAGTTCTGACCAGTGGAAGGGT 1387
          ||||||||||||||||| ||||||||||||| |||||||||||||||||||||||||||||
Db   6002 GATGATGCAGTCCTGTTACAAAGACGTTTGATATACAGCCAGTTCTGACCAGTGGAAGGGT 6061

Qy   1388 CGGAAAAAGTCTCAACATTAGGTCCCATTTGAAGCCAGTTCTGACCAGTCTGACGAAGGCGT 1447
          |||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||
Db   6062 CGGAAAAAGTCTCAACATTAGGTCCCATTTGAAGCCAGTTCTGACCAGTCTGACGAAGGCGT 6121

Qy   1448 CTGCACCTTTCTTCTGCGAGGAACTTGGTGTGGCTCACAGCTGAAAGATGAATGAATTAAGC 1507
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6122 CTGCACCTTTCTTCTGCGAGGAACTTGGTGTGGCTCACAGCTGAAAGATGAATGAATTAAGC 6181

Qy   1508 CGCCAGGCACCTATTGGAGGGACTTTCCAGCAGTTCAGAAGCAGCAGAACGATGTACATAGG 1567
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6182 CGCCAGGCACCTATTGGAGGGACTTTCCAGCAGTTCAGAAGCAGCAGAACGATGTACATAGG 6241

Qy   1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 1627
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   6242 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 6301

Qy   1628 CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGCCCAGA 1687
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   6302 CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGCCCAGA 6361

Qy   1688 GAGCTGCCTCCTGAGGAGAGAGAGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTTACGAAAGCAGGCTGAG 1747
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6362 GAGCTGCCTCCTGAGGAGAGAGAGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTTACGAAAGCAGGCTGAG 6421

Qy   1748 GAGGTCAATACTGAGTGGGGAAAAATTGAACCTGACTGGGCAGAGAGAAAAATA 1807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6422 GAGGTCAATACTGAGTGGGGAAAAATTGAACCTGACTGGGCAGAGAGAAAAATA 6481

Qy   1808 GATGAGACCCTGAAAGACTTCAGGAACTTCAAGAGGCCACGGATGCAGTCAGCTGGACCTCAAG 1867
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6482 GATGAGACCCTGAAAGACTTCAGGAACTTCAAGAGGCCACGGATGCAGTCAGCTGGACCTCAAG 6541

Qy   1868 CTGCGCCAAGCTGAGGTGATCAAGGGTCGGGGCCCGTGGGCCCGTGGGCCCGTCGGAGGAGATCCGCAGTCCATTGAC 1927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6542 CTGCGCCAAGCTGAGGTGATCAAGGGTCGGGGCCCGTGGGCCCGTGGGCCCGTCGGAGGAGATCCGCAGTCCATTGAC 6601

Qy   1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTGCCCTCTGAAA 1987
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   6602 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTGCCCTCTGAAA 6661

Qy   1988 GAGAACGTGAGCCACGTCAAGGGCACTTTGGGCCATTCAGCTC 2047
          |||||||||||||||||||||||||||||||||||||||||||
Db   6662 GAGAACGTGAGCCACGTCAAGGGCACTTTGGGCCATTCAGCTC 6721

Qy   2048 TCACCGTATAACCTCAGCACTCTGGAGAACACCAGACTGCAACACGAATGGAAGCTTCTGCAGGGTG 2107
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 7
AX538624
LOCUS　　AX538624　　11443 bp　　DNA　　linear　　PAT 23-NOV-2002
DEFINITION　Sequence 44 from Patent WO0229056.
ACCESSION　AX538624
VERSION　AX538624.1　GI:25271175
KEYWORDS　
SOURCE　synthetic construct
ORGANISM　synthetic construct
　　artificial sequences.
REFERENCE　1
AUTHORS　Chamberlain,J.S. and Harper,S.Q.
TITLE　Mini-dystrophin nucleic acid and peptide sequences
JOURNAL　Patent: WO 0229056-A 44 11-APR-2002;
　　THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES　Location/Qualifiers
source　　1..11443
　　/organism="synthetic construct"
　　/mol_type="unassigned DNA"
　　/db_xref="taxon:32630"
　　/note="Synthetic"
ORIGIN

Query Match　55.6%;　Score 1335.6;　DB 6;　Length 11443;

```
RESULT 8
AX538627
LOCUS       AX538627            12057 bp    DNA      linear   PAT 23-NOV-2002
DEFINITION  Sequence 47 from Patent WO0229056.
ACCESSION   AX538627
VERSION     AX538627.1  GI:25271181
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 47 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..12057
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

Query Match        55.6%;  Score 1335.6;  DB 6;  Length 12057;
Best Local Similarity 98.3%;  Pred. No. 0;
Matches 1350;  Conservative    0;  Mismatches   24;  Indels   0;  Gaps   0;
```

```
RESULT 9
HSDMDR
LOCUS       HSDMDR             12446 bp    mRNA     linear   PRI 12-SEP-1993
```

```
DEFINITION   Human mRNA for dystrophin.
ACCESSION    X14298
VERSION      X14298.1    GI:30845
KEYWORDS     Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 12446)
  AUTHORS    Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
             Davies,K.E.
  TITLE      Two human cDNA molecules coding for the Duchenne muscular dystrophy
             (DMD) locus are highly homologous
  JOURNAL    Nucleic. Acids Res. 17 (13), 5391 (1989)
  MEDLINE    89345106
   PUBMED    2668885
REFERENCE    2  (bases 1 to 12446)
  AUTHORS    Rosenthal,A.
  TITLE      Direct Submission
  JOURNAL    Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
             der DDr, Zentralinstitut fuer Molekularbiologie, Robert-Roessle
             Str.10, 1115 Berlin Buch, DDR
COMMENT      see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
             M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449,
             3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES             Location/Qualifiers
     source          1..12446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X chromosomal, Xp21."
                     /tissue_type="muscle"
                     /dev_stage="fetal and adult."
     CDS             99..11156
                     /note="unnamed protein product; dystrophin (AA 1 - 3685)"
                     /codon_start=1
                     /protein_id="CAA32479.1"
                     /db_xref="GI:30846"
                     /db_xref="GOA:P11532"
                     /db_xref="SWISS-PROT:P11532"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
```

```
FEATURES        Location/Qualifiers
     source     1..13957
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
                /note="EMBL/GenBank Accession No. M18533"

ORIGIN

Query Match            55.6%;  Score 1335.6;  DB 6;  Length 13957;
Best Local Similarity  98.3%;  Pred. No. 0;
Matches 1350;  Conservative 0;  Mismatches 24;  Indels 0;  Gaps 0;

Qy    1028 TGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGGTTCTTTACAGACACTCATAGA 1087
           ||||||||||||| |||||||||||||||||||||||||||||||||| ||||| |||||
Db    8216 TGGAGAAGCATTCATAAAAGGGTGAGTGAGACCTGGCAGAGGCTGCTTGGAGAAACTCATAGA 8275

Qy    1088 TTACTGCAACAGTTCCCCTGACCTGGAAAAGTTCTTGCTGGCTTACAGAGAGCTGAA 1147
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    8276 TTACTGCAACAGTTCCCCTGACCTGGAAAAGTTCTTGCTGGCTTACAGAGAGCTGAA 8335

Qy    1148 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGAAAAGGCTCCTAGAAGACTCCAAG 1207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8336 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGAAAAGGCTCCTAGAAGACTCCAAG 8395

Qy    1208 GGAGTAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGGTGAAATTGAAGCTCACACA 1267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8396 GGAGTAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGGTGAAATTGAAGCTCACACA 8455

Qy    1268 GATGTTTATCACAACCTGATGAAAAATCCTGAGATCCTGGAAGGTTCC 1327
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    8456 GATGTTTATCACAACCTGATGAAAAACAAAAATCCTGAGATCCTGGAAGGTTCC 8515

Qy    1328 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTCAAGTGAGTGGAACTT 1387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8516 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTCAAGTGAGTGGAACTT 8575

Qy    1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGCGT 1447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8576 CGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTGGAAGCGT 8635

Qy    1448 CTGCACCTTTCTCTGCCAGGAACTTCTCGGTGTGTGGCTACAGCTAGTGAAGATGAATTAAGC 1507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8636 CTGCACCTTTCTCTGCCAGGAACTTCTCGGTGTGTGGCTACAGCTAGTGAAGATGAATTAAGC 8695

Qy    1508 CGGCAGGCACCTATTGGAGGCGACTTTCAGAGTTCAGAAGCAGAGACGATCTACATAGG 1567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8696 CGGCAGGCACCTATTGGAGGCGACTTTCAGAGTTCAGAAGCAGAGACGATGTACATAGG 8755

Qy    1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 1627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8756 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 8815

Qy    1628 CGAATATTTCTGACAGAGGCCCTTGCTTGTTGGATCTCTTGAGCCTGTCA 1687
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8816 CGAATATTTCTGACAGAGGCCCTTGCTTGTTGGATCTCTTGAGCCTGTCA 8875

Qy    1688 GAGCTGCCTCCTGAGGAGAGAGAGCCCAGAAGTCACTCGGCTTCACCGAAAGCAGGCTGAG 1747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8876 GAGCTGCCTCCTGAGGAGAGAGAGCCCAGAAGTCACTCGGCTTCACCGAAAGCAGGCTGAG 8935

Qy    1748 GAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTTGCTGACTGGCAGAGAAAAATA 1807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8936 GAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTTGCTGACTGGCAGAGAAAAATA 8995

Qy    1808 GATGAGACCCTGAAAGACTTCAGGAACTTCAAGAGGGCACGGATGAGCTGACCTCAAG 1867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8996 GATGAGACCCTGAAAGACTTCAGGAACTTCAAGAGGGCACGGATGAGCTGACCTCAAG 9055

Qy    1868 CTGCGCCAAGCTGAGGTGATCAAGGGATCAAGGACTCCTGGCAGCCCGTGGGCGATCTCCTCATTGAC 1927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9056 CTGCGCCAAGCTGAGGTGATCAAGGGATCAAGGACTCCTGGCAGCCCGTGGGCGATCTCCTCATTGAC 9115

Qy    1928 TCTCTCCAAGATCACCTCGAGAAGTCAGGCACTTCGAGAGAAAATTGCGCCCTCTGAAA 1987
```

```
Qy    1568 GCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 1627
Db    8646 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 8705

Qy    1628 CGAATATTTCTGACAGAGCAGCAGCGCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGA 1687
           |||||||||||||||| ||| |||||||||||||||||||||||||||||||||||||
Db    8706 CGAATATTTCTGACAGAGCAGCAGCGCCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGA 8765

Qy    1688 GAGCTGCCTCCTGAGGAGAGAGAGCCCAGAAGTGCACTCTTGAGAAAAGCAGAGGCTGAG 1747
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    8766 GAGCTGCCTCCTGAGGAGAGAGAGCCCAGAAGTGCACTCTTGAGAAAAGCAGAGGCTGAG 8825

Qy    1748 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1807
           ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    8826 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8885

Qy    1808 GATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGAGGCCACGGATGAGCTGACCTTCAAG 1867
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    8886 GATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGAGGCCACGGATGAGCTGACCTTCAAG 8945

Qy    1868 CTGGCCCAAGCTGAGGTGATCAAGGGGATCAAGGACTCCTGGCAGCCCGTGGGCCGATCTCCTCATTGAC 1927
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8946 CTGGCCCAAGCTGAGGTGATCAAGGGGATCAAGGACTCCTGGCAGCCCGTGGGCCGATCTCCTCATTGAC 9005

Qy    1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGAGGAAAATTGCGCCCTCTGAAA 1987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9006 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGAGGAAAATTGCGCCCTCTGAAA 9065

Qy    1988 GAGAACGTGAGCCACGTCAATGACCTTACCACTTTTGGGCCATTCAGCTC 2047
           ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    9066 GAGAACGTGAGCCACGTCAATGACCTCTGCCAGCTTACCACTTTTGGGCCATTCAGCTC 9125

Qy    2048 TCACCGTATAAACCTCCAGCACTCTGGAAGACCTGAACACCAGGAGCTTCTGCGCAGGTG 2107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9126 TCACCGTATAAACCTCCAGCACTCTGGAAGACCTGAACACCAGGAGCTTCTGCGCAGGTG 9185

Qy    2108 GCCGTCGAGGACCGAGTCAGGCAGCGCGTGCGCAGCTGCATGAGAGCCCACAGGGACTTTGTCCAGCATCT 2167
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9186 GCCGTCGAGGACCGAGTCAGGCAGCGCGTGCGCAGCTGCATGAGAGCCCACAGGGACTTTGTCCAGCATCT 9245

Qy    2168 CAGCACTTTCTTTCCACGTCGTGTCCAGGTCCCTGGGGAGAGAGCCCATCTCGCCAAACAAA 2227
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9246 CAGCACTTTCTTTCCACGTCGTGTCCAGGTCCCTGGGGAGAGAGCCCATCTCGCCAAACAAA 9305

Qy    2228 GTGCCCTACTATATCAACCAGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACA 2287
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9306 GTGCCCTACTATATCAACCAGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACA 9365

Qy    2288 GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 2347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9366 GAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 9425

Qy    2348 ATGAAAACTCCGAAGACTCGCAGAGGCCCTTGCTTGTTGGATCTCTTGAGCCTGTCA 2401
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9426 ATGAAAACTCCGAAGACTCGCAGAGGCCCTTGCTTGTTGGATCTCTTGAGCCTGTCA 9479
```

RESULT 10
AX409637
LOCUS   AX409637   13957 bp   DNA   linear   PAT 14-JUN-2002
DEFINITION Sequence 2284 from Patent WO0229103.
ACCESSION   AX409637
VERSION   AX409637.1 GI:21442342
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE   Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2284 11-APR-2002;
            GENE LOGIC INC (US)

```
RESULT 11
AX538581
LOCUS       AX538581              13957 bp    DNA       linear   PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO0229056.
ACCESSION   AX538581
VERSION     AX538581.1  GI:25271086
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Chamberlain,J.S. and Harper,S.Q.
TITLE       Mini-dystrophin nucleic acid and peptide sequences
JOURNAL     Patent: WO 0229056-A 1 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match     55.6%;  Score 1335.6;  DB 6;  Length 13957;
Best Local Similarity   98.3%;  Pred. No. 0;
Matches 1350;  Conservative   0;  Mismatches  24;  Indels  0;  Gaps  0;
```

```
Qy   1028 TGGGCAAACATCTGTAGTGAGACAGAGACCGCTGGGTTCTTTACAAGACACTCATAGA 1087
          ||||||||||||||||| |       ||||||||| ||||||||||||||||||||||
Db   8216 TGGAGAAGCATTCATAAAGGGTGAGTGAGCGAGAGGCTGCTTTGGAAGAAACTTCATAGA 8275

Qy   1088 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCTGCTGCTTACCAGAAGCTGAA 1147
          ||||||||||||||||||| ||      |||||||||||||||||||||||||||||||
Db   8276 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCTGCTGCTTACCAGAAGCTGAA 8335

Qy   1148 ACAACTGCCAATGTCCTACACGGATGATGCTACCCGTAACGGAAAAGGCTCCAGAGAGACTCCCAAG 1207
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8336 ACAACTGCCAATGTCCTACACGGATGATGCTACCCGTAACGGAAAAGGCTCCAGAGAGACTCCCAAG 8395

Qy   1208 GGAGTAAAGAGAGCTGATGAAACAATGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 1267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9476 GAGCTCTTACCAGTTCTTAGCTGACCTGAATAATGTCAGAATTCAGATTTCTGAGCCTGCC 9535
```

```
Qy   2348 ATGAAACTCCGAAGACTGCGAAGAGGCCCTTGCTTGGATCCTCTTGAGCCTGTCA 2401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9536 ATGAAACTCCGAAGACTGCGAAGAGGCCCTTGCTTGGATCCTCTTGAGCCTGTCA 9589
```

```
LOCUS       HUMDYS               13957 bp    mRNA    linear    PRI 25-MAY-2000
DEFINITION  Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION   M18533 M17154 M18026 M20250
VERSION     M18533.1  GI:181856
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1699)
  AUTHORS   Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and
            Kunkel,L.M.
  TITLE     Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
            preliminary genomic organization of the DMD gene in normal and
            affected individuals
  JOURNAL   Cell 50 (3), 509-517 (1987)
  MEDLINE   87273512
  PUBMED    3607877
REFERENCE   2  (bases 1678 to 3830)
  AUTHORS   Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
  TITLE     Conservation of the Duchenne muscular dystrophy gene in mice and
            humans
  JOURNAL   Science 238 (4825), 347-350 (1987)
  MEDLINE   88018015
  PUBMED    3659917
REFERENCE   3  (bases 1 to 13957)
  AUTHORS   Koenig,M., Monaco,A.P. and Kunkel,L.M.
  TITLE     The complete sequence of dystrophin predicts a rod-shaped
            cytoskeletal protein
  JOURNAL   Cell 53 (2), 219-226 (1988)
  MEDLINE   88194521
  PUBMED    3282674
COMMENT     On May 25, 2000 this sequence version replaced gi:340693.
            Draft entry and computer-readable sequence kindly provided by
            M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
            determined by the size of the deleted DNA segment. Deletions found
            in different patients were from positions 302-2200, 473-1168,
            1691-1810, and 1169-3011.
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xp21.3-p21.1"
                     /tissue_type="muscle"
                     /dev_stage="fetus"
                     /note="G00-119-850"
     gene            1..13957
                     /gene="DMD"
     CDS             209..11266
                     /gene="DMD"
                     /codon_start=1
                     /product="dystrophin"
                     /protein_id="AAA53189.1"
                     /db_xref="GI:181857"
                     /translation="MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTQGKLPKEKGSRVHALNNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIIIHWQVKNVMKNIMAGLQQTNSEKILSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSIIMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEEHFQLHHQMHYSQ
                     QITVSLAQGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFPSQHLEAPEDKSFGSSL
                     MESEVNLDRYCTALEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH
                     QGRVGNILQLGSKLIGTGKLSEDEETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLKELNDWLTKTEERTRKMEEEPLGPDLEDLKRQVQDHKVLQEDLEQEQVRVN
```

ORIGIN

```
Query Match             55.6%;  Score 1335.6;  DB 9;  Length 13957;
Best Local Similarity   98.3%;  Pred. No. 0;
Matches 1350;           Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;
```

```
Qy   1028 TGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTCTTCTTTACAAGACACTCATAGA 1087
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   8216 TGGAGAAGCATTCATAAAGGGTGAGTGAGCCGCTGCTTGGAAGAAACTCATAGA 8275
```

```
Qy   1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCTGGCTTACAGAAGCTGAA 1147
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8276 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCTGGCTTACAGAAGCTGAA 8335
```

```
Qy   1148 ACAACTGCCAATGCTTCTACAGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 1207
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8336 ACAACTGCCAATGCTTCTACAGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 8395
```

Protein sequence (left column):
```
SLTHMVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWVLLQDILLKWQRLTE
EQCLFSAWLSEKEDAVNKIHTTGRQNEMLSSLQKLAVLKEDQNEMLSSLQKLAVLKEKKQSMGKLYSLK
QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTTQPSLTQTTVM
ETVTVTTREQIVKHAQELPPPPQKKRQITVDSEIRKRLDVDITELHSWITRSEA
VLQSPEFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSAQALVEQMVNEGVNA
DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLEQMTTAENWLKIQ
PTTPSEPTAIKSQLKICKDEVNRLSGLGQPQIERLKIQSIALKEKGQGPMFLDADFVAF
TNHFQVFSDVQAREKELQTIFDTLPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
EIMEQRLGELQALQSSLQEQQSGLYYLSTTVKEMSKKAPSEISRKYQSEFEEIEGRWNK
KLSSQLVENCQKLEBQMNKLRKIQNHIQTLKKWMAEVDVFLKEEWPALGDSBILKKQL
KQCRLLVSDIQTIQPSLNEVNEGGQKIKNEAEPFASRLETELKELNTWDWDHMCQQVY
ARKEALKGGLEKTVSLQKDLSEMHEWMTQAEEBYLERDFEYKTPDELQKAVEMKRAK
EEAQQKEAKVKLLIESVNSVIAQAPPVAQEALKKELETLTTNVQWLCRLNGKCKTLE
EVWACWHELLSYLEKANKWLNEVFKLKTENIPGGAEISEVIDSLENLMRHSEDNP
NQIRILAQTLTDGGVMDELINEBLETFNSRWRELHEBAVRRQKLLEQSIQSAQETEKS
LHLIQESLTFIDKQLAAYIADKVDAAQMPEQAQKIQSDLTSHEISLEEMKKHNQGKEA
AQRVLSQIDVAQKLQDVSMKFRLFQKPANFELRLQESKMILDEVKMHLPALETKSVE
QEVVQSOLNHCVNLYKSLSEVKSEVEMVIKTGRQIVQKKQTENPKEELDERVTALKLHY
NELGAKVTERKQQLEKCLKLSRKMREMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
VAWGKATQKEIEKQKVHLKSITEVGEALKTVLGKKETLVEDKLSLLNSNWIAVTSRAE
EWLNLLLEYQKHMETFDQNVDHITKWIIQADTLLDESEKKKPQQKEDVLKRLKAELND
IRPKVDSTRDQAANLMANRGDHCRLVEPQISELNHRFAAISRIKTKASIPLEGKOLKE
QFNSDIQKLLEPLEAEIQQGVNLKEEDPNKDMNEDNEGTVKELLQRGDNLQQRITDER
KREEIKIKQQLLQTKHNALKDLRSQRRKKALEISHOWTQYKRQQADDLLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKKEELNAVRRQAEGLSEDGAAMAVEPTIQLSKRWRE
IESKFAQFRRLNFAQIHTVREETMWMTEDMPLEISVYPSTYLTEITHVSQALLEVEQ
LLNAPDLCAKDFEDLFKQESLKNIKDSLQQSSGRIDIIHSKKTAALQSATPVBRVKL
QEALSQLDFQEVNKMYKDRQGRFPDRSVBKMRFRFHYDIKIFNQWLTEAEOFLRKTQI
PENWEHAKYKWYLKELQDCIGQRQTVRTLNATGEEIIQQSKTDASILQEKLGSLNL
RWQEVCKQLSDRKKRLEEQKNILSEFQRDLNEFVLWLEEADNIASIPLEPGKEQOLKE
KLEQVLLVBELPLRQGILKQLNETGGPLVSAPISPEQDKLENKLKQTNLQIKVS
RALPEKQGEIEAQIKDLQLEKKLEDLEEQLNHLLLWLSPIRNQLEIYNQPNOEGPFD
VQEITEIAVQAKQPDVEEILSKGQHLYKEKPATQPVKRKLEDLSSEWKAVNRLLQELRA
KQPDLAPGLTTIGASPTQPVTLVTQPVVKETAISKLEMPSSLMLEVPALADPNRAWT
ELTDWLSLLDQVIKSQRVMVGDLEDINEMIIKQKATMODLEQRRPQLEELITAAQNLK
NKTSNQEARTIITDRIERIQNQWDEVQDNLQNRRQQLNEMLKDSTQWLEAKEABQVL
GQARAKLESWKEGPTVDAIEAHTDVHNLDENSOKILRSLEGSD
DTRKVHMITENINASWRSIHKRVSEREAALEETHRLLQQFPLDLEKFLAWLIEABTTA
NVLQDATRKERLLEDSKGVKELMKWQDLQGEIEAHTDVYHNLDENSOKILRSLEGSD
DAVLIQRRLDNMNFKWSELRKKSLNIRSHLEASSDQWKRIHLSLQELLVWLQKDDEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLYQ
EPRELPEERAQVNTWLRLKQABEVNTEWELNLHSADWQRKIDETLERLQELQEATD
ELDLKRQAEVIKGSWQPVGDLIJDSLQDHLEKVKALRGEIAPLKENVSHVNDLARQL
TTLGIQLSPYNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVQGP
WERAISPNKVPYYINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA
LCLDLLSLSAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHNNLVNVPLCVD
MCLNWLLNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKRYLFKQVASSTGFCDQRRL
GLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQPANNKPEIPAALFLDWMRLEPQSM
VWLPVLHRVAABETAKHQAKCNICKECPIIGFRYRSLKFNYDICQSCFFSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLSGDNMET
PVTLJNFWYDSAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSYLNDSISPNESID
DEHLLIQHYCQSLNQDSPLSQPRSPAQLISLESEBRGELERILADLEEENRNLQAEY
DRLKQQHEHKGLSPIJSPSPBEMMTSPSQSPRDAELIAEAKLLRQHKGRLEARMQILEDH
NKQLESQLHRLRQLLEQPQAEAKYNGTTVSSPSTSLQRSDSSQPMLLRVVGSQTSDSM
GEEDLLSPPQDTSTGLEBVMEQLNNSFPSSRGRNTPGKPMREDTM"
```

```
Qy   1208  GGAGTAAAAGAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA  1267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8396  GGAGTAAAAGAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA  8455

Qy   1268  GATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCC  1327
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8456  GATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCC  8515

Qy   1328  GATGAGGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT  1387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8516  GATGAGGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT  8575

Qy   1388  CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT  1447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8576  CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT  8635

Qy   1448  CTGCACCTTTCTGCAGGAACTTCTGCAGCTGAAAGATGATGAATTAAGC  1507
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   8636  CTGCACCTTTCTGCAGGAACTTCTGCAGCTGAAAGATGATGAATTAAGC  8695

Qy   1508  CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCCAGAAGCAGATGCTGTACATAGG  1567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8696  CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCCAGAAGCAGATGCTGTACATAGG  8755

Qy   1568  GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA  1627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8756  GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA  8815

Qy   1628  CGAATATTTCTGACAGAGCAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGA  1687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8816  CGAATATTTCTGACAGAGCAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGA  8875

Qy   1688  GAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCACCAAAGCAGGCTGAG  1747
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8876  GAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCACCAAAGCAGGCTGAG  8935

Qy   1748  GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGCAGAGAAAAATA  1807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8936  GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGCAGAGAAAAATA  8995

Qy   1808  GATGAGACCCTTGAAAGACTTCAGGAACTTCAGAGGGCCACGGACTGAGCTGACCTTCAAG  1867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8996  GATGAGACCCTTGAAAGACTTCAGGAACTTCAGAGGGCCACGGACTGAGCTGACCTTCAAG  9055

Qy   1868  CTGCGCCAAGCTGAGGTGATCAAGGATCAAGGGATCAGGGCAGCCCGTGGCAGCCCTCATTGAC  1927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9056  CTGCGCCAAGCTGAGGTGATCAAGGATCAAGGGATCAGGGCAGCCCGTGGCAGCCCTCATTGAC  9115

Qy   1928  TCTCTCCAAGATCACCTGAGAAAGTCACTCAAGGCACTTCAGGAGAAAGTTGCGCCTCTGAAA  1987
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9116  TCTCTCCAAGATCACCTGAGAAAGTCACTCAAGGCACTTCAGGAGAAAGTTGCGCCTCTGAAA  9175

Qy   1988  GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACACTTTGGGCATTCAGCTC  2047
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9176  GAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACACTTTGGGCATTCAGCTC  9235

Qy   2048  TCACCGTATAACCTCAGCACTTCAGCAGCTCAGAGTCGAGGACCGAGTCAGGGCCACAGAGATTAAGC  2107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9236  TCACCGTATAACCTCAGCACTTCAGCAGCTCAGAGTCGAGGACCGAGTCAGGGCCACAGAGGTG  9295

Qy   2108  GCCGTCGAGGACCGAGTCAGGGCCACCTATTGGAGGCCACAGGGACTTTGGTCCAGCATCT  2167
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9296  GCCGTCGAGGACCGAGTCAGGGCCACCTATTGGAGGCCACAGGGACTTTGGTCCAGCATCT  9355

Qy   2168  CAGCACTTTCTTTCCACGTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA  2227
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9356  CAGCACTTTCTTTCCACGTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA  9415

Qy   2228  GTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACA  2287
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9416  GTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACA  9475

Qy   2288  GAGCTCTACCAGTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTCAGATTCCAGA  2347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   9476  GAGCTCTACCAGTCTTAGCTGACCTGAATAATGTGCTTCAGTCAGCTGAATAATGTTCAGCTTCAGCTTGAGCCTTCA  9535

Qy   2348  ATGAAACTCCGAAGACTGCAGAGAGGCCCTTGCTTGGCTTCCTGGATCCTGGATCTCTGAGCCTGTCA  2401
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9536  ATGAAACTCCGAAGACTGCAGAGAGGCCCTTGCTTGGATCCTGGATCTCTGAGCCTGTCA  9589
```

```
RESULT 13
AR220819
LOCUS       AR220819                13977 bp    DNA        linear    PAT  26-SEP-2002
DEFINITION  Sequence 60 from patent US 6426186.
ACCESSION   AR220819
VERSION     AR220819.1    GI:23327696
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1  (bases 1 to 13977)
AUTHORS     Jones,K.A., Volkmuth,W. and Walker,M.G.
TITLE       Bone remodeling genes
JOURNAL     Patent: US 6426186-A 60 30-JUL-2002;
FEATURES             Location/Qualifiers
source       1..13977
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match    55.6%;  Score 1335.6;  DB 6;  Length 13977;
Best Local Similarity 99.3%;  Pred. NO. 0;
Matches 1350;  Conservative   0;  Mismatches   0;  Indels    24;  Gaps   0;
```

```
Qy   1028  TGGGCAAACATCTGAGATGGACAGAGACCGCTGGGTTCTTTACAAGACACTCATAGA  1087
           ||| | || ||||||||||||||||||||||||||||||||||||||||||||||||
Db   8216  TGGAGAAGCATTCATAAAAGGTGAGTGAGCGAGGGTGAGTGAGCTGCTTTGGAGAAACTCATAGA  8275

Qy   1088  TTACTGCAACAGTTCCCCTGACCTGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAA  1147
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8276  TTACTGCAACAGTTCCCCTGACCTGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAA  8335

Qy   1148  ACAACTGCCAATGTCTCACAGGATGCTACCCCGTAAGAGAAGGCTCCTAGAGAAGGTCCTAGAGACTCCAAG  1207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8336  ACAACTGCCAATGTCTCACAGGATGCTACCCCGTAAGAGAAGGCTCCTAGAGACTCCAAG  8395

Qy   1208  GGAGTAAAAGAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA  1267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8396  GGAGTAAAAGAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA  8455

Qy   1268  GATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCC  1327
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8456  GATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCTGGAAGGTTCC  8515

Qy   1328  GATGAGGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT  1387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8516  GATGAGGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT  8575

Qy   1388  CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT  1447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8576  CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT  8635

Qy   1448  CTGCACCTTTCTGCAGGAACTTCTGCAGCTGAAAGATGATGAATTAAGC  1507
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   8636  CTGCACCTTTCTGCAGGAACTTCTGCAGCTGAAAGATGATGAATTAAGC  8695

Qy   1508  CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCCAGAAGCAGATGCTGTACATAGG  1567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8696  CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCCAGAAGCAGATGCTGTACATAGG  8755

Qy   1568  GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA  1627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8756  GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA  8815

Qy   1628  CGAATATTTCTGACAGAGCAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGA  1687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8816  CGAATATTTCTGACAGAGCAGCAGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCCAGA  8875
```

```
2b  8816 CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGA 8875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  1688 GAGCTGCCTCCTGAGGAGAGCCAGCAGCCTTTCACTCGGCTTCTACGAAAGCAGCCTGAG 1747

2b  8876 GAGGTCAATACTGAGTGAGGAAAAATTGAACCTGCACTCCGCTTGACTGGCAGAGAAAATA 8935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  1748 GAGGTCAATACTGAGTGAGGAAAAATTGAACCTGCACTCCGCTTGACTGGCAGAGAAAATA 1807

2b  8936 GATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGATGAGCTGGACCTCAAG 8995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  1808 GATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGATGAGCTGGACCTCAAG 1867

2b  8996 GATGAGACCCTTGAAAGACTTCCAGGATCAAGGGATCCTGCAGCCCGTGGGCAGCCCAAG 9055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  1868 CTGGCCAAGCTGAGGTGATCAAGGGATCCTGCAGCCCGTGGGCAGCCCAAG 1927

2b  9056 CTGGCCAAGCTGAGGTGATCAAGGGATCAAGGGATCCTGCAGCCCGTGGGCCGATCCTCTCATTGAC 9115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  1928 TCTCTCCAAGATCACCTCGAGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCCTCTGAAA 1987

2b  9116 TCTCTCCAAGATCACCTCGAGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCCTCTGAAA 9175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  1988 GAGAACGTGAGCCACGCAGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTC 2047

2b  9176 GAGAACGTGAGCCACGCAGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTC 9235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  2048 TCACCGTATAAACCTTCAGCACTCTCCGGAAGACTCTGAACACCAGATGGAAGCTTCTCGCAGGTG 2107

2b  9236 TCACCGTATAAACCTTCAGCACTCTCCGGAAGACTCTGAACACCAGATGGAAGCTTCTCGCAGGTG 9295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  2108 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCT 2167

2b  9296 GCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCT 9355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  2168 CAGCAGACTTCTCTTCCACGTCTGTCCGGAGGAGAGCCATCCGCCAAACAAATGACA 2227

2b  9356 CAGCAGACTTCTCTTCCACGTCTGTCCGGAGGAGAGCCATCCGCCAAACAAATGACA 9415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  2228 GTGCCCTACTATATCAACCACCAGACTTCTGGAATAAGTCAGATTCTCAGCCAAAATGACA 2287

2b  9416 GTGCCCTACTATATCAACCACCAGACTTCTGGAATAAGTCAGATTCTCAGCCAAAATGACA 9475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  2288 GAGCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 2347

2b  9476 GAGCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 9535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
2y  2348 ATGAAAACTCCGAGAACTCCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA 2401

2b  9536 ATGAAAACTCCGAGAACTCCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCA 9589
```

/db_xref="taxon:9615"
<1..289
290..11332
/codon_start=1
/product="dystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLWWEVEDCYEREDVQKTFTKVWNAQFSKFGKQHIENLFSDL
QDGRRLLDLLEGLTGQKLPKEKGSRTVHALNNNVNKALRVLQKNNVDIVNIGSTDIVDG
NHKLTLGLIWNIILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIAKYQLGIEKLLDPEDV
ATTYPDKKSIMYITSLFQVLPQQVSIEAIQEVEMLPRPSQVTREEHFQIHHQMHYSQ
QITVSLAQGYERAPSFPKPRFKSYAYTQAAYVVTSDPTRSPLPSQHLETPEDKSFGRS
LTETEANLDSYQTALEEVLSWLLSAEDALQAQGEISNDVEEVKEQFHTHEGYMMDLTS
HQGRVGNVLQGSQLIGTGKLSEDETEVQEQEMLLNSRWECLRVASMEKQSNLHKVL
MDLQNQQLKELNDWLTKTEERTRMKEPLGPDIEDLKRQVQQHKVLQEDLEQEQVRV
NSLTHMVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWVLLQDILLKWQRFT
EEQCLFSAWLSEKEDAVNKIHTTGFKDQSEVLSNLQKLAVLKTDLEKKKQTMDKLCSL
NQDLLSALKNTVVAHKMEAWLDNFAQRWDNLVQKLEKSSAQISQAVVTTQPSILQTTV
METVMVTTREHILVKHAQEELPPPPQKRQIIVDSEIRKRLDVDITELHSWITRSE
AVLQSPEFAIYRKEGNFSDLKEKVNAIEREKAEKPRKLQDASRSAQALVEQMVNEGVN
ADSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIITFYNQLQQLEQMTTAENWLKT
QPTTTSEPTAIKSQLKICKDEINRLSALDQPIERLKIQSIALKEKGQGPMFLDADFVA
FTNHENQVFADVQAREKELQTIFDSLPMRYQETMSTILTWIQQSETKLSIPQVTVTE
YDIMEQRLGELQALQSSLQDQQNGLNYLSTTVKEMSKKALPSDISRKYQSEFEIEGR
WKKLSSQLVEHCQKLEEQMAKLRKIQNHIKTLKKWITEVDVFLKEEWPALGDSEILKR
QLKQCRLLVNDIQIIQPSLNSVNEGAQKMKNEAEPEFAGRLETELRLENTQWDYMCRQ
VYARKEALKGGLDKTVSLQKDLSEMHEWMTQAEEEYLERDFEYKTDELQTAVEEMKR
AKEEAQEAKVKLLTESVNSVIAQPPAAQEALKELDTLTTVYQWLCTRLNGKCKT
LEEWACWHELLSYLEKANKWLSEVEVKLKTTENISGGAEEIAEVLDSLENLMQHSED
NPNQIRILAQTLITDGGVMDELINELETNSRWRELHEAVRQKLLEQSIQSAQEIE
KSLHLIQESLSSIDKQLAAYIADKVDAAQMPEQAQKIQSDLTSHEISLEEMKKHNQGK
ETAQRVLSQIDVAQKKLQDVSMKFRLFQKPANFEQRLQESKMILDEVKMILPALETKS
VEQEVVQSQLNHCVNLYKSLSEVKSEVENVIKTGRQIVQKKQTENPKEDERVTALKL
HYNELGAKVTERKQQLEKCLKLSRKMRKEENALTEWLAATDMELTKRSAVEGMPSNLD
SEVAWGKATQKEIEKQKVHLKSVTEVGEALKTVLGKKEMLVEDKLSLNSNWIAVTSR
AEEWLNLLLEYQKHMETFDQNVDYITNWIQADALLDESEKKPQQKEDILKRLKAEM
NDIRPKVDSTRDQAANLMANRGDHCRKVVEPKISELNHRFAAISHRIKTGKASIPLKE
LEQFNSDIQKLLEPLEAEIQQGVNLKEEDFNKDMSEDNEGTVKELLQRGDNLQQRITD
ERKREIKIKQQLLQTKHNALKDLRSQRRKKALEISHQWYQYKRQADDLLKCLDDIEK
KLASLPERDERKIKEIDRELQKKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRW
REIESKFAQFRRLNFAQIHIVHEESVVAVTEDMPLEISYVPSTYLTEITHVSQALSEV
EELLNAPDLCAQDFPEDLFKQEESLKNIKDSLQQISGRIDIIHNKKTAALHSATPAERA
KLQEALSRLDFQWERVNNMYKDRQGRFDRSVEKWRRFHYDMKILNQWLTEAEQFLKKT
QIPENWEHAKYKWYLKELQDGIGQRQSVVRVLNATGEEIIQQSSKTDASILQEKLGSL
NLRWQEVCKQLAERKRLEEQKNILSEFQRDVNEFVLWLEEADNVANIPEFQNEQQL
KEKLEQVKLLAEELPLRQGILKQLNETGGTVLVSAPLSPEEQDKLENKLKQTNLQWIK
VSRNLPEKQEEIEAHVKDLGQLEEQLNHLLLWLSPIRNQLEIYNQPNQTGPFDIKEIE
VAVQAKQPDVEEGILSKGQHIYKPATQPAKRKLEDLSSDWKVTVQLIQELRAKQPGP
APGLTTVRAPPSQVTVLVTQPAVTKETAISKLMPSSLLLEVPALADFNRAWTELTDW
LSLLDDRVIKSQRVMVGDLEDINEMIIKQKATLQDLEQRRPQLELITAAQNLKNKTSN
QEARTIITDRIERIQSQWDEVQEHLQNRRLQLITMLKDSTQWLEAKEEAEQVLGQARA
KLESWKEAPYTVDAIQKKITETKQLAKDLRQWQINVDVANDLALKLLRDYSADDTRKV
HMITENINASWASIHKRLSEREAALEETHRLLQQFPLDIEKFLAWLTEAETTANVJQD
ATHKRLLEDSKGVRELMKWQWDLQGEIEAHTDIYHNLDENGQVLRSLEGSDDAALL
QRRLDMNFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELVWLQLKDDELSRQAP
IGGDPPAVQKQNDVHRAFKRELKTKSPVIMSTLETVRIFLTEQPLEGLEKLYQEPREL
PPEERAQNVTRLLRKQAEVNTQWEKLNVHSADMQRKIDEALERLQELQEATDELDLK
LRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEITPLKENVSYVNDLARQLTTLGI
QLSPYNLNTLEDLNTRWKLLQVAIEDRIRQLHEAHRDFGPASQHLEQEHNLNVPLCVDMCLNW
SPNKVPYINETQTTCWDHPKMTLLYQSALDLNNVRFSAYRTAMKLRRLQALCLDL
LLNVDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKAVASSTGFCQQRRLGLLLH
DSIQIPRQLGEVASFGGSNIEPSVRSCFQPANNKPEIEAALFLDWRLEPQSMWLPV
LHRVAAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDICQSCFFSGRVAKGHKMHYP
MVEYCTPTTSGEDVRDFAKVLKNKFRTKRYFAKHPRMGYLPVQTVLEGDNMETPVTLI
NFWPVDSAPASSPQLSHDDTHSRIEHYASRLKMENSNGSYLNDSISPNESIDDEHLL
IQHYNRSLNQESPLSQPRSPAQIHLSLESEERGELERILADLEGRNRNLQAEYDRLKQ
QHEHKGLSPLPSPBEMMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDHNKQE
SQLHRLRQLLEQPQAEAKVNGTTVSSPSTSLQRSDSSQPMLLRVVGQTSESMGEEDL
LSPPQDTSTGLEEVMEQLNHSFPSSRGRNTPGKPMREDTM"
11333..>13887
11789..11828
```

  5' UTR
  CDS

  3' UTR
  repeat_region

ORIGIN
                /rpt_type=tandem
                /rpt_unit="ca"

Query Match   49.1%;   Score 1179.4; DB 4; Length 13887;
Best Local Similarity   91.2%; Pred. No. 4.8e-293;
Matches 1252; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

Qy 1028 TGGGCAAACATCTGTAGATGGACAGAAGACCGTGGTTCTTTACAAGACACTCATAGA 1087
          ||||||||||||||||||||||||||||| ||||||| |||||||||||||||||||
Db 8282 TGGGCAAGCATCCATAAAGATTGAGTGAGTGAGGACTTGCTCTGGAAGAAACCACAGA 8341

Qy 1088 TTACTGCAACAGTTCCCCCTGGACCTGGAAAGTTTCTTGCTGCTGGCTTACAGAAGCTGAA 1147
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db 8342 TTACTGCAACAGTTCCTGCTGGACCTGGCTTACAGAAGCTGGCTTACAGAAGCCGAA 8401

Qy 1148 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 1207
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 8402 ACAACTGCCAACGTCCTGCAGGATGCTCGAGGATGAAAGGCTTCAGAAGAGTTCTCAAG 8461

Qy 1208 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGTGAAATTGAAGCTCACACA 1267
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db 8462 GGAGTAAGAGAGCTGATGAAACAATGGCGAAAATGGCCAAAAGTCCCAAGGAGGAAGGTTCT 8521

Qy 1268 GATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCTGAGATCCCTGGAGGTTCC 1327
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8522 GATATCTATCACAACCTGGACCGACGAAAATGGCAAATCCTGAGATCTCTTGGAGGTTCT 8581

Qy 1328 GATGATGCAGTCCTGTTGACAAAAGACGTTTGGATAACACATGAACTTCAAGTGGAGTGAACTT 1387
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 8582 GACGATGCAGCCTTGTGTGCAAGACACGTTTGGATAACACATTCAAGTGGAGGGAACTT 8641

Qy 1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGGGT 1447
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db 8642 CGGAAAAAGTCTCTCAACATTAGGTCTCACTTGGAAGCCAGTTCTGACCAGTGGAAGCGT 8701

Qy 1448 CTGCACCTTTCTCTGCGAGGGAACTTCTGGTGTGGCTCACAGCTGAAAGATGATGAATTAAGC 1507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8702 CTGCACCTTTCTCTTCAGGGAACTTCTGGATGGTGGCTCCCAGCTGAAAGATGAGTTAAGC 8761

Qy 1508 CGGCAGGCACCTATTGGAGCCGACTTTCAGCAGTTCAGAAGCAGAAGCAGATGTACATAGG 1567
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db 8762 CGGCAGGCACCCATTGGAGGAGACTTTCCAGCGGTCGAGAGCAGAGCAGAATGATGTACACAGG 8821

Qy 1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 1627
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8822 GCCTTCAAGAGGGAATTGAAAACGAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA 8881

Qy 1628 CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAACTCTACCAGGAGCCCAGA 1687
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db 8882 CGAATATTTCTGACAGAGCAGCCTTTAGAAGGACTAGAGAAACTCTACCAGGAGCCCAGA 8941

Qy 1688 GAGCTGCCTCCTGAGAGAGAGCAGCCCAGAATGTCACTCGGCTTCACGAAAGAGAAGCGTGAG 1747
          ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db 8942 GAGCTGCCTCCTGAGAGAGAGCAGCCCAGAATGTCACACGGCTCCACGAAAGAGAAGCGTGAG 9001

Qy 1748 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGACTCCCGCTGACTGGCAGAGAGAAAATA 1807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9002 GAGGTCAACACTCAGTGGGAAAAAACTGAACGTGACTCTCGAGACTGGCAGAGAGAAAAATA 9061

Qy 1808 GATGAGACCCCTGAAAGACTTCAAGAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1867
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 9062 GACGAGGCCCCTGAAAGACTTCAGGAGACTCCAGGAGCAACAGATGAGCTGGAAGACTCCAAG 9121

Qy 1868 CTGCGCCAAGCTGAGGTGACCAAGGTGATCAAAGGATCAAAGGATGCTCGGGCCGCAGCCCGGT 1927
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db 9122 CTACGTCAGGAGGTGACCACCTTGACCAAAGGATGACCGGTGTGGGTGACCTCCTCATTGAC 9181

Qy 1928 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCAAGCCACTTGCGCCTCTGAAA 1987
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9182 TCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTACACCTCTGAAA 9241

Qy 1988 GAGAACGTGAGCCACGTCACAATGCACCTTGCTCGCCAGGCTTACCACTTTGGGCATTCAGCTC 2047
          |||| ||||||||||||||| |||||||||||||||||||||||||||||||||||||||

Db 9242 GAGAATGTCAGCTACGTCAATGACCTTGCTCGGCTGACAGCTTGCTCCGGAAGACACCAGAGAAGCTGAACACCACCAGCAGATGGGAAGCTTGGGCCGTTCTGCAGGTG 9301
          |||| |||| ||| |||| |||||||||||| |||||||||| ||||| || ||||||||||||||||||||||||||||||||||
Qy 2048 TCACCGTATAAACCTCAGCACTCTGGAAGACCTGAAGACCTGAACACCACCAGCAGATGGGAAGCTTGGGCCGTTCTGCAGGTG 2107
          ||||||||| |||||||||||||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db 9302 TCACCATATAAACCTCAACACTCTGGAAGACCTGAAGACCTGAACACCACCAGCAGATGGGAAGCTTGGGCCGTTCTGCAGGTG 9361

Qy 2108 GCCGTGCAGGACCGGAGTCAGGCAGCTGCATGAGAGCGTCGATGAGAGCCCACAGGGACTTTGGTCCAGCATCT 2167
          |||| |||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db 9362 GCCATTGAGGACCGGCATCAGAGCTGCATGAGAGCGTCGATGAGAGCCCACAGGGACTTTGGACCAGCCTCC 9421

Qy 2168 CAGACTTCTTCTTCCACGTCTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGGCCAAACAAA 2227
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 9422 CAGCACTTCCTTCCTTCGTCTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCACCAAACAAA 9481

Qy 2228 GTGCCCTACTATATCAACCACGAGACTCAACACTTGCTGGGACCATCCCAAAATGACA 2287
          ||||||||||||| |||||||||||||| |||||||||||||| ||||||||||||||
Db 9482 GTGCCCTACTATATCAACCACGAGACCCCAAAACTTGCTGGGACCATCCCAAAATATGACA 9541

Qy 2288 GAGCTCTACCAGTCTTTAGCTGACCTGACCTGATAATGTCAGATTCTCAGATTCTTGAGCCTTC 2347
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db 9542 GGCTCTCCACCAGTCTTTAGTGACCTGATAATGTCAGTTCTTCAGCTTCTCAGCTTACAGGACTGCC 9601

Qy 2348 ATGAAACTCCGAAGACTGCAGAAGACTGCAGAAGGCCCTTGCTTGGATCCTGGAATCTTTC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9602 ATGAAACTCCGAAGACTGCAGAAGACTGCGAAGGCCCTTTGCTTGGATCTCTTGAGGCCTATC 9654

RESULT 15
AX306153
LOCUS   AX306153   13815 bp   DNA   linear   PAT 11-DEC-2001
DEFINITION   Sequence 904 from Patent WO0188188.
ACCESSION   AX306153
VERSION   AX306153.1   GI:17645441
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
  TITLE   Method for examining ischemic conditions
  JOURNAL   Patent: WO 0188188-A 904 22-NOV-2001;
                School Juridical Person Nihon University (JP)
FEATURES   Location/Qualifiers
     source   1..13815
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

Query Match   47.6%;   Score 1143.6; DB 6; Length 13815;
Best Local Similarity   89.5%; Pred. No. 8.3e-284;
Matches 1230; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 1028 TGGGCAAACATCTGTAGATGGACAGAAGACCGTGGTTCTTTTACAAGACACTCATAGA 1087
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||| |||
Db 8207 TGGGGAAACATTCATAAAGAGTAGTGAGTAGTGAGGACTAAGCAAGAAGGCTGCTTTTGGAAGAAAACTCATAGA 8266

Qy 1088 TTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCTGCCTGGCTTACAGAAGCTGAA 1147
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 8267 TTACTGCAGCAGTTCCTGCTGGACCTGGCTTACAGAAGTTTCTTTCCTGATTACGGAAGCAGAA 8326

Qy 1148 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 1207
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db 8327 ACAACTGCCAATGTCCTGCAGGATGCCTGACGCTTCCCGTAAGAAGTCCTAGAAGACTCCAGG 8386

Qy 1208 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGTGAAATTGAAGCTCACACA 1267
          ||||||||||||||||||||||||| |||||||||||||||||||| |||||||||||
Db 8387 GGAGTCAGGAGCTGATGAAACCATGGCAAGATCCCAAGGAGAGATCCCAAGAGCAAGAATTGAAACTCACACA 8446

Qy 1268 GATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCGATCCCTGGAGGTTCC 1327
          || ||| |||||||||||||||||||||||||||||| |||||||||||||||

```
Db    8447  GATATCTATCACAATCTTGATGAAAATGGCCAAAAAATCCTGAGATCCCTGGAAGGTTCG  8506

Qy    1328  GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT  1387
                  |||||| ||||||||||| || ||||||||||||||||||||| ||||||||||||
Db    8507  GATGAAGCACCCCTGTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAACTT  8566

Qy    1388  CGGAAAAAGTCTCTCAACATTAGGTCCATTTGGAGCCAGTTCTGACCAGTGGAAGCGT  1447
                 ||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||
Db    8567  CAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCAAGTTCTGACCAGTGGAAGCGT  8626

Qy    1448  CTGGCACCTTCTCTGCAGGAACTTCTCGTGTCGCTACAGCTGAAAGATGATGAATTAAGC  1507
             |||||  |||| |||||  ||||| || |||  || |||||||  |||||||||||| |||
Db    8627  TTGCATCTTTCTTCAGGAACTTCTTGTTTGGCTACAGCTGAAAGATGATGAACTGAGC  8686

Qy    1508  CGGCAGCCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGG  1567
             |||||||||||| ||||||||| |||||||||||||| |||||||||||| |||||||||
Db    8687  CGTCAGGCACCCCTGGTGATTTCCCAGCAGTTCAGAAGCAGAATGATATACATAGG  8746

Qy    1568  GCCTTCAAGAGGGAATTGAAAACTAAAGAACTGTAATCATGAGTACTCTTGAGACTGTA  1627
             |||||||||||||||||||||||||||||||| ||||||||||||| ||||||||||| |
Db    8747  GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTCGGAGACTGTG  8806

Qy    1628  CGAATATTTTCTGACAGCAGCAGCCTTTGGAAGGACTAGAGAAACTCACCAGGAGCCCAGA  1687
             |||| ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    8807  AGAATATTTCTGACAGGAGCAGCCTTTGGAAGGACTAGAGAAACTCACCAGGAGCCCAGA  8866

Qy    1688  GAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCACTGAAAGCAGGCTGAG  1747
             ||||||||||||||||||||||||| || ||||| |||||||||||||||||||||||
Db    8867  GAACTGCCTCCTGAAGAGAGCTCAGAATGTCACTCGGCTCCTACGAAAGCAGCCTGAA  8926

Qy    1748  GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA  1807
             ||||||||  || ||| || |||||||||| |||||||| ||||| ||||||||||||||
Db    8927  GAGGTCAACGCTGAATGGCACAAATTGAACCTGCGCGCTCAGCTGATTGGCAGAGAAAAATA  8986

Qy    1808  GATGAGAATCGAGTGATCAAGAACTTCAAGAGGCCACGGATGAGCTGGACCTCCAAG  1867
             ||||||  |||||||||||||||||||||||||||||||||||||| |||||||||||
Db    8987  GATGAAGCTCTTGAAAGATCCAGGAATCTCAGGAAGCTGGCAGCTGGACCTCCAAG  9046

Qy    1868  CTGCGCGCCAAGTGATGAGTGATCAAGGACATCCTGGCCAGCCCCGGGGCGATCCTCATTGAC  1927
             ||||||||| ||||||| ||||||||||||||||||||| ||||||||||||||||||||||
Db    9047  TTGCGCGCCAAGTGATGAGTGATCAAGGACATCCTGGCCAGCCCGGGGCGATCCTCATTGAC  9106

Qy    1928  TCTCTCCAAGATCACCTGAGAAGTCAAGGACTTCGAGAAATTGCGCCTTCGAAA  1987
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    9107  TCTCTGCAAGATCACCTTGAAAGTCAAGGACACTTCGGGGAGAAAATTGCACCTTCTTAAA  9166

Qy    1988  GAGAACGTGAGCCACGTCAATGACTCTGCCAGCTTACCACTTTGGGCATTCAGCTC  2047
             |||||| ||| ||||| |||||||| ||||||||||||||| |||||||||||||
Db    9167  GAGAATGTCAATGCGTCAATGACCACTTGCACATCAGCTGACCACACTGGGCATTCAGCTC  9226

Qy    2048  TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTG  2107
             |||||||||||| ||||||||| |||||||||||||||| |||| |||||||||||||
Db    9227  TCACCTTATAACCTCAGCACTTTGGAAGATTGAATACCAGATTGAGCCTTCACAGGTG  9286

Qy    2108  GCCGTGCGAGGACCGAGTCAGGCAGTCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCT  2167
             |||||| |||||||| ||||| |||||||||||| ||||||| |||||||||||||||||||
Db    9287  GCTGTGGGAGGACCGTGTCAGACAGCTGCTGATGAAGCCCACAGGGACTTTGTCTCCAGCATCC  9346

Qy    2168  CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA  2227
             ||||||||| ||||| ||||||||||||||| |||||||||||||||||| |||||||||
Db    9347  CAGCACTTCCTTTCCACTTCAGTTCAGGGTCCCTGGGAGAGAGCCATCTCACCAAAACAAA  9406

Qy    2228  GTGCCCTACTATATCAACCACAGAGACTCAAACAACTGCTGGGACCATCCCAAAATGACA  2287
             ||||| ||||||||||||||||||||||| |||||| ||| ||||||| |||||||||||
Db    9407  GTGCCCTACTATATCAACCACACCACTTGTTGGGACCACCCCAAAATGACA  9466

Qy    2288  GAGCTCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC  2347
             ||||||||| ||||| ||||| |||||||||||||  |||| ||||||| ||||||||||
Db    9467  GAGCTCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGGTTCTCCGCGTATAGGACTGCC  9526

Qy    2348  ATGAAACTCCGAAGACTGCAGAAGGCGAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA  2401
             |||||||| |||||||||||||| |||||||||||||||||| ||||||||||||||||||
Db    9527  ATGAAGCTCAGAGACTGCAGAAGGCGTCCAGAAGGCCCTTTGCTTGCTTGGATCTCTTGAGCCTGTCA  9580
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: April 4, 2004, 11:55:32 ; Search time 619.552 Seconds
                                     (without alignments)
                                     16463.377 Million cell updates/sec

Title: US-09-845-416-9_COPY_600_3000
Perfect score: 2401
Sequence: 1 adatgcattcaacatcgcca.........tggatctcttgagcctgtca 2401

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
              1: geneseqn1980s:*
              2: geneseqn1990s:*
              3: geneseqn2000s:*
              4: geneseqn2001as:*
              5: geneseqn2001bs:*
              6: geneseqn2002s:*
              7: geneseqn2003as:*
              8: geneseqn2003bs:*
              9: geneseqn2003cs:*
             10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2401	100.0	3858	6	AAD37237	Aad37237 Human dys
2	2401	100.0	4825	6	AAD37257	Aad37257 Adeno-ass
3	2401	100.0	4848	6	AAD37263	Aad37263 Adeno-ass
4	2401	100.0	5060	6	AAD37264	Aad37264 Adeno-ass
5	2250	93.7	3999	6	AAD37234	Aad37234 Human dys
6	2250	93.7	4966	6	AAD37256	Aad37256 Adeno-ass
7	2250	93.7	4990	6	AAD37262	Aad37262 Adeno-ass
8	2067	86.1	4182	6	AAD37230	Aad37230 Human dys
9	2067	86.1	5149	6	AAD37255	Aad37255 Adeno-ass
10	1839.4	76.6	5462	6	ABK81999	Abk81999 DNA encod
11	1737	72.3	3531	6	AAD37238	Aad37238 Human dys
12	1737	72.3	4498	6	AAD37258	Aad37258 Adeno-ass
13	1725	71.8	5339	6	ABK81998	Abk81998 DNA encod
14	1695	70.6	3510	6	AAD37240	Aad37240 Human dys
15	1695	70.6	4476	6	AAD37259	Aad37259 Adeno-ass
16	1419	59.1	5417	5	ABK81997	Abk81997 DNA encod
17	1335.6	55.6	5952	6	AAD06794	Aad06794 Human dys
18	1335.6	55.6	8689	6	ABK82000	Abk82000 DNA encod
19	1335.6	55.6	11058	6	AAD37229	Aad37229 Human dys
20	1335.6	55.6	11241	6	ABK82005	Abk82005 cDNA enco
21	1335.6	55.6	11443	6	ABK82002	Abk82002 DNA encod
22	1335.6	55.6	12923	1	AAN90338	Aan90338 Sequence
23	1335.6	55.6	13957	6	ABK81959	Abk81959 cDNA enco
24	1335.6	55.6	13957	6	ABT10904	Abt10904 Human bre
25	1335.6	55.6	13957	6	ABN95786	Abn95786 Gene #228
26	1335.6	55.6	13957	6	ABS69900	Abs69900 Human dys
27	1335.6	55.6	13977	6	ABS70403	Abs70403 Human bon
28	1323	55.1	2169	6	AAD37232	Aad37232 Human dys
29	1263	52.6	4414	6	AAD37260	Aad37260 Adeno-ass
30	1251	52.1	3446	6	AAD37242	Aad37242 Human dys
31	1143.6	47.6	13815	6	AAV18885	Aav18885 Mus muscu
32	1143.6	47.6	13815	6	ABK81960	Abk81960 cDNA enco
33	1143.6	47.6	13815	6	ABI99799	Abi99799 Mouse isc
34	1143.6	47.6	19307	2	AAT27558	Aat27558 Shuttle v
35	1093.8	45.5	4402	3	AAZ48568	Aaz48568 A rod sho
36	1092.2	45.5	1991	6	AAD37231	Aad37231 Human dys
37	1077	44.9	1667	6	AAD37235	Aad37235 Human dys
38	1021.2	42.5	4402	6	AAZ48567	Aaz48567 A rod sho
39	1018.8	42.4	4075	3	AAZ48569	Aaz48569 A rod sho
40	976	40.6	1821	6	AAD37241	Aad37241 Human dys
41	750	31.2	1340	6	AAD37239	Aad37239 Human dys
42	709.2	29.5	3747	3	AAZ48566	Aaz48566 A rod sho
43	633.2	26.4	3275	1	AAN97129	Aan97129 Partial s
44	589	24.5	1434	6	AAD37243	Aad37243 Human dys
45	498.6	20.8	10705	7	ABT41896	Abt41896 Toxicity

ALIGNMENTS

RESULT 1
AAD37237
ID   AAD37237 standard; DNA; 3858 BP.
XX
AC   AAD37237;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO20018695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 48-49; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

```
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;

Query Match            100.0%;  Score 2401;  DB 6;  Length 3858;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 2401;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy       1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA  659

Qy      61 TGTTGATACCACCTACCTTCCAGATAAGAAGTCCATCTTAATGTACATGTCACATCCTCCA  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 TGTTGATACCACCTACCTTCCAGATAAGAAGTCCATCTTAATGTACATGTCACATCCTCCA  719

Qy     121 AGTTTGCCTCAACAAGTGAGCATTGAGAGCCATCCAGGAAGTTGCCAAGGCC          180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 AGTTTGCCTCAACAAGTGAGCATTGAGAGCCATCCAGGAAGTTGCCAAGGCC          779

Qy     181 ACCTAAAGTGACTAAAGAAGAAGAACATTTCAGTACATCATCAAATGCACTATTCTCAACA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 ACCTAAAGTGACTAAAGAAGAAGAACATTTCAGTACATCATCAAATGCACTATTCTCAACA  839

Qy     241 GATCACGGTCAGTCTAGCACAGGGATATGAGAAGAACTTCTCCCCTAAGCCTGATTCAA    300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 GATCACGGTCAGTCTAGCACAGGGATATGAGAAGAACTTCTCCCCTAAGCCTGATTCAA    899

Qy     301 GAGCTATGCCTACACACAGCAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCGCCCATT  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 GAGCTATGCCTACACACAGCAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCCATT  959

Qy     361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAAGCTTCAGTTGATGGAGAG       420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     960 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCAAGCTTCAGTTGATGGAGAG      1019

Qy     421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAGTATTATCGTGGCTTCTTTC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAGTATTATCGTGGCTTCTTTC 1079

Qy     481 TGCTGAGGACACATTGCAAGCACAAGGACAAGGAGAGATTCTAATGATGTGGAAGTGGAAAGA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1080 TGCTGAGGACACATTGCAAGCACAAGGACAAGGAGAGATTCTAATGATGTGGAAGTGGAAAGA 1139

Qy     541 CCAGTTCATATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGTTGG     600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1140 CCAGTTCATATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCGGTTGG    1199

Qy     601 TAATATTCTACAATTGGGAGGTAAGCTGATTGGACCAGAAAATTATCAGAAGATGAAGA   660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1200 TAATATTCTACAATTGGGAGGTAAGCTGATTGGACCAGAAAATTATCAGAAGATGAAGA  1259

Qy     661 AACTGAAGTACACAGAGAGCAGATGAATCCTCCTAAATTCAAGATGGAATGGCTCAGGGTAGC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1260 AACTGAAGTACACAGAGAGCAGATGAATCCTCCTAAATTCAAGATGGAATGGCTCAGGGTAGC 1319

Qy     721 TAGCATGCTGAAAAACAAACAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1320 TAGCATGCTGAAAAACAAACAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1379

Qy     781 GAAAGAGTTGAATGACTGGCTAACAAAACAGGAGAAAATGGAGGAGA               840
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    1380 GAAAGAGTTGAATGACTGGCTAACAAAACAGGAGAAAATGGAGGAGA              1439

Qy     841 GCCTCTTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1440 GCCTCTTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 1499

Qy     901 AGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1500 AGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTGGTAGT 1559

Qy     961 TGATGAATCTAGTGGAGATCACCGCAACTGCTGTTTGGAAGAACAACTTAAGGTATTGGG   1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1560 TGATGAATCTAGTGGAGATCACCGCAACTGCTGTTTGGAAGAACAACTTAAGGTATTGGG   1619

Qy    1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAGACCGCTGGGTTCTTTACAAGACAC   1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1620 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAGACCGCTGGGTTCTTTACAAGACAC   1679

Qy    1081 TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA   1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1680 TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA   1739

Qy    1141 AGCTGAAACAACTGCCAATGTCCTACAGATGCCAATGTCCGTAAGGAAAGGCTCCTAGAGA  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1740 AGCTGAAACAACTGCCAATGTCCTACAGATGCCAATGTCCGTAAGGAAAGGCTCCTAGAGA  1799

Qy    1201 CTCCAAGGGAGTAAAAGAGGTGATGAAACAATGGAAATGTTGCCAAGGCC          1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1800 CTCCAAGGGAGTAAAAGAGGTGATGAAACAATGGAAATGTTGCCAAGGCC          1859

Qy    1261 TCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCTGGA   1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1860 TCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAAATCCTGAGATCCTGGA   1919

Qy    1321 AGGTTCCGATGATGGCAGTCCTGTGTACAAAGACGTTTGGATAAACAACGTTCAAGTGGAG  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1920 AGGTTCCGATGATGGCAGTCCTGTGTACAAAGACGTTTGGATAAACAACGTTCAAGTGGAG  1979

Qy    1381 TGAACTTCGGAAAAGTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG   1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1980 TGAACTTCGGAAAAGTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG   2039

Qy    1441 GAAGGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2040 GAAGGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA  2099

Qy    1501 ATTAAGCCGGAGCCGGCACCCATTGGAGGCGCACCCATTGGAGGCGCACGATGT      1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2100 ATTAAGCCGGAGCCGGCACCCATTGGAGGCGCACCCATTGGAGGCGCACGATGT      2159

Qy    1561 ACATAGGGCCTTCAAGAGCGAATTGAAAATCTAAAGAACCTAAGAACTGAGTCATGAGAACTCTTGA 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2160 ACATAGGGCCTTCAAGAGCGAATTGAAAATCTAAAGAACCTAAGAACTGAGTCATGAGAACTCTTGA 2219

Qy    1621 GACTGTACAGAATATTTCTGACAGAGCCTTTGGAAGACTAGAGAAACTCTACCAGGA   1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2220 GACTGTACAGAATATTTCTGACAGAGCCTTTGGAAGACTAGAGAAACTCTACCAGGA   2279

Qy    1681 GCCCAGAGAGTCGCCTCCTGAGGGACCCAGAATGTCACTCGGCTTCTACGAAAGCA   1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2280 GCCCAGAGAGTCGCCTCCTGAGGGACCCAGAATGTCACTCGGCTTCTACGAAAGCA   2339

Qy    1741 GGCTGAGGAGTCAATACTGAGTGGGGAAAATTGAACCTGACTGGCCAGAG         1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2340 GGCTGAGGAGTCAATACTGAGTGGGGAAAATTGAACCTGACTGGCCAGAG         2399

Qy    1801 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTCAGAACT               1860
           |||||||||||||||||||||||||||||||||||||||||||
Db    2400 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTCAGAACT               2459

Qy    1861 CCTCAAGCTGCGCGCCAAGCTGATCAAGGGTCATCAAGGGATCTCCT           1920
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    2460 CCTCAAGCTGCGCGCCAAGCTGATCAAGGGTCATCAAGGGATCTCCT           2519

Qy    1921 CATTGACTCTCTCAAGATCCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2520 CATTGACTCTCTCAAGATCCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC  2579

Qy    1981 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTACCACTTTGGGCAT 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2580 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGTTACCACTTTGGGCAT 2639
```

CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;

Query Match   100.0%; Score 2401; DB 6; Length 4825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAAACTACTGATCCTGAAGA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1357 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAAACTACTGATCCTGAAGA 1416

Qy    61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACACTCTTCCA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1417 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACACTCTTCCA 1476

Qy   121 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1477 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1536

Qy   181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACATTCTCAACA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1537 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACATTCTCAACA 1596

Qy   241 GATCACGGTCAGTCTAGCACACAGGGATATGAGATATGAGAAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1597 GATCACGGTCAGTCTAGCACACAGGGATATGAGATATGAGAAGAACTTCTTCCCCTAAGCCTCGATTCAA 1656

Qy   301 GAGCTATGCTACACACAGCCTGCTTATGTCACCACCTTCACCACTCTGACCCTACACGGAGCCCCATT 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1657 GAGCTATGCTACACACAGCCTGCTTATGTCACCACCTTCACCACTCTGACCCTACACGGAGCCCCATT 1716

Qy   361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1776

Qy   421 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1777 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

Qy   481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGTGTGAAGTGGTGAAAGA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGTGTGAAGTGGTGAAAGA 1896

Qy   541 CCAGTTTCATACTTCATGAGGGGTACATGATGATTGACAGCCATCAGGGCCGGGTTGG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1897 CCAGTTTCATACTTCATGAGGGGTACATGATGATTGACAGCCATCAGGGCCGGGTTGG 1956

Qy   601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2016

Qy   661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGGCTCAGGGTAGC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2017 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGGCTCAGGGTAGC 2076

Qy   721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCGAATCAGAAACT 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCGAATCAGAAACT 2136

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGGAAAGAAGGAAATGGAGGAAGA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGGAAAGAAGGAAATGGAGGAAGA 2196

Qy   841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGTGCTTCA 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  2041 TCAGCTCCACGTATAACCTCAGCACTCTGAGAAGACCAGATGGAAGCTTCT 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2640 TCAGCTCCACGTATAACCTCAGCACTCTGGAAGACCAGATGGAAGCTTCT 2699

Qy  2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCAGCTGCATGAAGCCCACCAGGGACTTTGGTCC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2700 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCAGCTGCATGAAGCCCACCAGGGACTTTGGTCC 2759

Qy  2161 AGCATCTCAGCACTTCTTCTTCCACGTCTCGTCCAGGGTCCCTGGGAGGAGAGCCATCTCGCC 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2760 AGCATCTCAGCACTTTCTTTCTTCCACGTCTCGTCCAGGGTCCCTGGGAGGAGAGCCATCTCGCC 2819

Qy  2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTGCTCGGGACCATCCCAA 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2820 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTGCTCGGGACCATCCCAA 2879

Qy  2281 AATGACAGAGCTCTTACCAGTCTTTAGCTGACCTGACCTGAATAAATGTCAGATTCTTCAGCTTATAG 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2880 AATGACAGAGCTCTTACCAGTCTTTAGCTGACCTGACCTGAATAAATGTCAGATTCTTCAGCTTATAG 2939

Qy  2341 GACTGCCATGAAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTTCTTGAGCCTGTC 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2940 GACTGCCATGAAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTTCTTGAGCCTGTC 2999

Qy  2401 A 2401
         |
Db  3000 A 3000
```

RESULT 2
AAD37257

ID   AAD37257 standard; DNA; 4825 BP.
XX
AC   AAD37257;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 61-62; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The

Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTCAAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTTGGTGGTGGTAGT 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 2316

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTCGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTCGCTGCTTGGAAGAACAACTTAAGGTATTGGG 2376

Qy 1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2377 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAC 2436

Qy 1081 TCATAGAATTACTGCAACAGTTCCCCTGGACCTGGAAAGTTCTTGCCTGGCTTACAGA 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2437 TCATAGAATTACTGCAACAGTTCCCCTGGACCTGGAAAGTTCTTGCCTGGCTTACAGA 2496

Qy 1141 AGCTGAAACAACAACTGCCAATGTCCTACAGATGTCCTACACCCGTAAGGAAAGGCTCCTAGAAGA 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2497 AGCTGAAACAACAACTGCCAATGTCCTACAGATGTCCTACACCCGTAAGGAAAGGCTCCTAGAAGA 2556

Qy 1201 CTCCAAGGGAGTAAAAGAGTCTGATGAAAACAATGGCAAGACCTCCAAGTGAAATTGAAGC 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2557 CTCCAAGGGAGTAAAAGAGTCTGATGAAAACAATGGCAAGACCTCCAAGTGAAATTGAAGC 2616

Qy 1261 TCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGA 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2617 TCACACAGATGTTTATCACAACCTGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGA 2676

Qy 1321 AGGTTCGATGATGCAGTCTCTGTTACAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2677 AGGTTCGATGATGCAGTCTCTGTTACAAGACGTTTGGATAACATGAACTTCAAGTGGAG 2736

Qy 1381 TGAACTTCGGAAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTG 1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2737 TGAACTTCGGAAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTG 2796

Qy 1441 GAAGCGTCTGCACCTTTCTCTGCGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2797 GAAGCGTCTGCACCTTTCTCTGCGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2856

Qy 1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGCACTTTCCAGCAGTTCAGAGAGCAGAACGATGT 1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2857 ATTAAGCCGGCAGGCACCTATTGGAGGCGCACTTTCCAGCAGTTCAGAGAGCAGAACGATGT 2916

Qy 1561 ACATAGGGCCTTCAAGGATGAATTGAAAACTAAAGAACCTGTAATCATGATGACTACTCTTGA 1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2917 ACATAGGGCCTTCAAGGATGAATTGAAAACTAAAGAACCTGTAATCATGATGACTACTCTTGA 2976

Qy 1621 GACTGTACGAATATTTCTGACACAGAGCAGCCGTTGGAAGACTAGAGAAACTCTACCAGGA 1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2977 GACTGTACGAATATTTCTGACACAGAGCAGCCGTTGGAAGACTAGAGAAACTCTACCAGGA 3036

Qy 1681 GCCCAGGAGAGCTGCCTCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3037 GCCCAGGAGAGCTGCCTCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 3096

Qy 1741 GGCTGAGGAGGTCAATACTGAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCCAGAG 1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3097 GGCTGAGGAGGTCAATACTGAGTGGAAAAATTGAACCTGCACTCCGCTGACTGGCCAGAG 3156

Qy 1801 AAAAATAGATGAGACCCTTGAAAGAACTCAAGAGGCCACGGATGAGCTGA 1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3157 AAAAATAGATGAGACCCTTGAAAGAACTCAAGAGGCCACGGATGAGCTGA 3216

Qy 1861 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTCTGCAGCCCGTGGGCGATCTCCT 1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3217 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTCTGCAGCCCGTGGGCGATCTCCT 3276

Qy 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3277 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 3336

Qy 1981 TCTGAAAAGAGAACCTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCACTTGGGCAT 2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3337 TCTGAAAGAGAACCTGAGCCACGTCAATGACCTTGCTGCCAGCTTACCACTTGGGCAT 3396

Qy 2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACCACCAGATGGAAGCTTCT 2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3397 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACCACCAGATGGAAGCTTCT 3456

Qy 2101 GCAGGTGGCCGTCGAGGACCGAGTCGAGTCGCATGGAAGCCACAGGGGACTTTGGTCC 2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3457 GCAGGTGGCCGTCGAGGACCGAGTCGAGTCGCATGGAAGCCCACAGGGGACTTTGGTCC 3516

Qy 2161 AGCATCTCAGCACTTCTTTCCACGTCGTCCACGGGTCCCTGGGAGGAGAGCCATCTCGCC 2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3517 AGCATCTCAGCACTTCTTTCCACGTCGTCCACGGGTCCCTGGGAGGAGAGCCATCTCGCC 3576

Qy 2221 AAACAAAGTGCCCTACTATATCAACACAGAGCTCAAACAACTTGCTGGGACCATCCCAA 2280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3577 AAACAAAGTGCCCTACTATATCAACACAGAGCTCAAACAACTTGCTGGGACCATCCCAA 3636

Qy 2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3637 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3696

Qy 2341 GACTGCCATGAAACTCCGAGAACTGCAGAAGGCCCTTTGCTTGGATCCTTGCAGTCGTC 2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3697 GACTGCCATGAAACTCCGAGAACTGCAGAAGGCCCTTTGCTTGGATCCTTGCAGTCGTC 3756

Qy 2401 A 2401
           |
Db 3757 A 3757

RESULT 3
AAD37263
ID   AAD37263 standard; DNA; 4848 BP.
XX
AC   AAD37263;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Cytomegalovirus.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 68-70; 71pp; English.

```
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a
CC   cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ   Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 2401;  DB 6;  Length 4848;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2401;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 1439

Qy         61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1499

Qy        121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1559

Qy        181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 1619

Qy        241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1679

Qy        301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1739

Qy        361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1799

Qy        421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1859

Qy        481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1860 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1919

Qy        541 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1979

Qy        601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2039

Qy        661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2040 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2099

Qy        721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2100 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2159

Qy        781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2160 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA 2219

Qy        841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 2279

Qy        901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2280 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 2339

Qy        961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2340 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2399

Qy       1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2400 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 2459

Qy       1081 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2460 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 2519

Qy       1141 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2520 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 2579

Qy       1201 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2580 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2639

Qy       1261 TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2640 TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 2699

Qy       1321 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2700 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 2759

Qy       1381 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2760 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 2819

Qy       1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2820 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2879

Qy       1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2880 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 2939

Qy       1561 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2940 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2999

Qy       1621 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3000 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 3059

Qy       1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3060 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 3119

Qy       1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3120 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 3179

Qy       1801 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3180 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 3239

Qy       1861 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGGCGATCTCCT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 70-71; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a
CC   small polyA signal sequence
XX
SQ   Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

Query Match   100.0%; Score 2401; DB 6; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACATGCATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAACTACTGATCCTGAAGA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1592 ACATGCATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAACTACTGATCCTGAAGA   1651

Qy   61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCATCATCACTCTTCCA   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1652 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCATCATCACTCTTCCA   1711

Qy   121 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1712 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   1771

Qy   181 ACCTAAAGTGACTAAAGAAGAAGAACATTTTCAGTTACATCATCAAATGACTAATTCTCAACA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1772 ACCTAAAGTGACTAAAGAAGAAGAACATTTTCAGTTACATCATCAAATGACTAATTCTCAACA   1831

Qy   241 GATCACGGTCAGTCTAGCACAGGGATATGCACAGCAGGATATGCACAGCAGAACTTCTTCCCCTAAGCCTCGATTCAA   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1832 GATCACGGTCAGTCTAGCACAGGGATATGCACAGCAGGATATGCACAGCAGAACTTCTTCCCCTAAGCCTCGATTCAA   1891

Qy   301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1892 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT   1951

Qy   361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1952 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   2011

Qy   421 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTCTAATGATGTGGAAGTGTGAAAGA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2012 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTCTAATGATGTGGAAGTGTGAAAGA   2071

Qy   481 TGCTGAGGAGCACAATTGCAAGCACAGAGAGAGATTCTAATGATGTGGAAGTGTGAAAGA   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2072 TGCTGAGGAGCACAATTGCAAGCACAGAGAGAGATTCTAATGATGTGGAAGTGTGAAAGA   2131

Qy   541 CCAGTTTCATACTTCATGAGGGGTACATGATTGACAGCCCATCAGGGGGCCGGGTTGG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2132 CCAGTTTCATACTTCATGAGGGGTACATGATTGACAGCCCATCAGGGGGCCGGGTTGG   2191

Qy   601 TAATATTCTACAATTGGAACTGAAGCTAAGCTCCTAAATTCAAGATGGAAGATGAAGA   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2192 TAATATTCTACAATTGGAACTGAAGCTAAGCTCCTAAATTCAAGATGGAAGATGAAGA   2251

Qy   661 AACTGAAGTACAAGAGCAGATGAATCCTAAATTCAAGATGGAATGGCCTCAGGGTAGC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2252 AACTGAAGTACAAGAGCAGATGAATCCTAAATTCAAGATGGAATGGCCTCAGGGTAGC   2311

Db   3240 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGAATCCTGGCAGCCCCGTGGGCGATCCT   3299

Qy   1921 CATTGACTTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAGAAATTGCGCC   1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3300 CATTGACTTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAGAAATTGCGCC   3359

Qy   1981 TCAGCTCTCACGGTATAACCTCAGACTCTGGAAGACCTGCGCCAGCTTACCACTTTGGGCAT   2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3360 TCAGCTCTCACGGTATAACCTCAGACTCTGGAAGACCTGCGCCAGCTTACCACTTTGGGCAT   3419

Qy   2041 TCAGCTCTCACGGTATAACCTCAGACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT   2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3420 TCAGCTCTCACGGTATAACCTCAGACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT   3479

Qy   2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCC   2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3480 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCC   3539

Qy   2161 AGCATCTCAGCACTTCTTCCACGTCTCGTCTGCTCCAGGGTCCCTGGGAGAGGAGCCATCTCGCC   2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3540 AGCATCTCAGCACTTCTTCCACGTCTCGTCTGCTCCAGGGTCCCTGGGAGAGGAGCCATCTCGCC   3599

Qy   2221 AAACAAAGTGCCCACTACTATAATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA   2280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3600 AAACAAAGTGCCCACTACTATAATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA   3659

Qy   2281 AATGACAGAGCTCTACCAGTCTTAGCTGACCTGAATAAATGTCAGATTCTCAGCTTATAG   2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3660 AATGACAGAGCTCTACCAGTCTTAGCTGACCTGAATAAATGTCAGATTCTCAGCTTATAG   3719

Qy   2341 GACTGCCATGAAACTCGAGAACTCGAGAAGGCCAGAGACTCGCAGAGGCCCTTGCTTGGATCTCTTGAGCCTGTC   2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3720 GACTGCCATGAAACTCCGAGAACTCGAGAAGGCCAGAGACTGCAGAGGCCCTTGCTTGGATCTCTTGAGCCTGTC   3779

Qy   2401 A 2401

Db   3780 A 3780

RESULT 4
AAD37264
ID   AAD37264 standard; DNA; 5060 BP.
XX
AC   AAD37264;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Cytomegalovirus.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular

Qy 1801 AAAAATAGATGAGACCCTTGAAGACTTCAAGAGGCCACGGATGAGCTGGA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3392 AAAAATAGATGAGACCCTTGAAGACTTCAAGAGGCCACGGATGAGCTGGA 3451

Qy 1861 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3452 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 3511

Qy 1921 CATTGACTCTCTCCAAGATCACCTGAGAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3512 CATTGACTCTCTCCAAGATCACCTGAGAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 3571

Qy 1981 TCTGAAAGAGAACGTGAGCGACGTTCACCGCCAGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3572 TCTGAAAGAGAACGTGAGCGACGTTCACCGCCAGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 3631

Qy 2041 TCAGCTCTCACGTATAACCTTCAGCACTCTGGAGGACCTTGGGAGCTTCT 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3632 TCAGCTCTCACGTATAACCTTCAGCACTCTGGAGGACCTTGGGAGCTTCT 3691

Qy 2101 GCAGGTGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAGAAGCCCACAGGGACTTTGGTCC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3692 GCAGGTGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAGAAGCCCACAGGGACTTTGGTCC 3751

Qy 2161 AGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGTGCCCTGGGAGGAGAGCCATCTCCGCC 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3752 AGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGTGCCCTGGGAGGAGAGCCATCTCCGCC 3811

Qy 2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3812 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA 3871

Qy 2281 AATGACAGAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3872 AATGACAGAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3931

Qy 2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3932 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 3991

Qy 2401 A 2401
         |
Db 3992 A 3992

RESULT 5
AAD37234
ID   AAD37234 standard; DNA; 3999 BP.
XX
AC   AAD37234;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX

Qy   721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCCCAGAATCAGAAACT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2371

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGGAAAAATGGGAGGAAGA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGGAAAAATGGGAGGAAGA 2431

Qy   841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2432 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 2491

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTTGGTAGT 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2492 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGGTTGGTAGT 2551

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAGAACAACTTAAGGTATTGGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2552 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAGAACAACTTAAGGTATTGGG 2611

Qy 1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2612 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 2671

Qy 1081 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTCTTGCTGGCTTACAGA 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2672 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTCTTGCTGGCTTACAGA 2731

Qy 1141 AGCTGAAACAACTGCCAATGTCCTACAGGATGTCTACCCGTAAGGAAAGGTCCTAGAAGA 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2732 AGCTGAAACAACTGCCAATGTCCTACAGGATGTCTACCCGTAAGGAAAGGTCCTAGAAGA 2791

Qy 1201 CTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2792 CTCCAAGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2851

Qy 1261 TCACAACAGATGTTATCACAACCTGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGGA 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2852 TCACAACAGATGTTATCACAACCTGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGGA 2911

Qy 1321 AGGTTCCCATGATGATGCAGTCCTGTTACAAGAGACGTTTGGATAACATTCAAGTTGGAG 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2912 AGGTTCCCATGATGATGCAGTCCTGTTACAAGAGACGTTTGGATAACATTCAAGTGGGAG 2971

Qy 1381 TGAACTTCGGAAAAAGTCTCTCTCAACATTAGGTCCCATTTGGAAGCGCCAGTTCTGACCAGTG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2972 TGAACTTCGGAAAAAGTCTCTCTCAACATTAGGTCCCATTTGGAAGCGCCAGTTCTGACCAGTG 3031

Qy 1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3032 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 3091

Qy 1501 ATTAAGCCGGCACCAGCACCTATTGGAGGCGCACTTCCAGCAGTTCAGAAGCAGAACGATGT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3092 ATTAAGCCGGCACCAGCACCTATTGGAGGCGCACTTCCAGCAGTTCAGAAGCAGAACGATGT 3151

Qy 1561 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTAGTACTTCTTGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3152 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTAGTACTTCTTGA 3211

Qy 1621 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGGACTAGAGAAACTTCTACCAGGA 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3212 GACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAGGACTAGAGAAACTTCTACCAGGA 3271

Qy 1681 GCCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAGCCCAGAATGTCACTCGGCTTCTACGGAAAGCA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3272 GCCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAGCCCAGAATGTCACTCGGCTTCTACGGAAAGCA 3331

Qy 1741 GGCTGAGGAGGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTCCGCTCCGCACTCGGCAGAG 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3332 GGCTGAGGAGGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTCCGCTCCGCACTCGGCAGAG 3391

```
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 46-47; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;

Query Match       93.7%;  Score 2250;  DB 6;  Length 3999;
Best Local Similarity   94.5%;  Pred. No.0;
Matches 2401;   Conservative   0;   Mismatches   0;   Indels 141;   Gaps   1;
```

```
Qy     1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 659

Qy    61 TGTGATACCACCTATCCAGATAGAAGTCCATCTTAATGTACATCACCATCACTCCTTCCA 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660 TGTGATACCACCTATCCAGATAGAAGTCCATCTTAATGTACATCACCATCACTCCTTCCA 719

Qy   121 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAGTGAAATGTGCCAAGGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAGTGAAATGTGCCAAGGCC 779

Qy   181 ACCTAAAGTGACTAAAGAAGAGAACATTTCAGTTACATCACAAATGCACTATTCCAACA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780 ACCTAAAGTGACTAAAGAAGAGAACATTTCAGTTACATCACAAATGCACTATTCCAACA 839

Qy   241 GATCACGGTCAGTCTAGCACAGGGATATGAGACAGGAACTTCTCCCCTAAGCCTGATTCAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   840 GATCACGGTCAGTCTAGCACAGGGATATGAGACAGGAACTTCTCCCCTAAGCCTGATTCAA 899

Qy   301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCACGGAGCCCATT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCACGGAGCCCATT 959

Qy   361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGAGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 TGAAGTAAACCTGGACCGTTGATACCCGGTTATCAAAGACAGCTTTAGCGTTCTTTC 1079

Qy   421 TGAAGTAAACCTGGACCGTTGATACCCGGTTATCAAAGACAGCTTTAGCGTTCTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 TGAAGTAAACCTGGACCGTTGATACCCGGTTATCAAAGACAGCTTTAGCGTTCTTTC 1079

Qy   481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy   541 CCAGTTTCATACTCATGAGGGTTACATGATGGACCTGACAGCCCATCAGGGCCGGGTTGG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1140 CCAGTTTCATACTCATGAGGGGTACATGATGGACCTGACAGCCCATCAGGGCCGGGTTGG 1199

Qy   601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259
```

```
Qy   661 AACTGAGTACAAGAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260 AACTGAGTACAAGAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy   721 TAGCATGGAAAACAAAGCAATTACATAGAGTTTAATGATCGTCCAGAATCAGAAACT 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1320 TAGCATGGAAAACAAAGCAATTACATAGAGTTTAATGATCGTCCAGAATCAGAAACT 1379

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAACAAGGAAAATGGAGGAAGA 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1380 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAACAAGGAAAATGGAGGAAGA 1439

Qy   841 GCCTCCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1440 GCCTCCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 1499

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACTCACATGGTGGTGTAGT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1500 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACTCACATGGTGGTAGT 1559

Qy   961 TGATGAATCTAGTGGAGATCACGGCAACTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1560 TGATGAATCTAGTGGAGATCACGCAACTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619

Qy  1021 AGATCGATGGGCAAACTCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAGAC-- 1078
           |||||||||||||||||| |||||||||||||||||||||||||||||| ||||||
Db  1620 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTTACAGACCA 1679

Qy  1079 ------------------------------------------------------------ 1079
Db  1680 GCCTGACCTAGCTCCTGACTGACCACTATTGGAGCCTCTCTACTCAGACTGTACTCT 1739

Qy  1079 ------------------------------------------------------------ 1079
Db  1740 GGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC 1799

Qy  1079 --------ACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAA 1119
                   |||||||||||||||||||||||||||||||||||||||
Db  1800 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAA 1859

Qy  1120 GTTTCTTGCCTGCTTACAGAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCG 1179
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1860 GTTTCTTGCCTGCTTACAGAGAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCG 1919

Qy  1180 TAAGGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAGCTGTGATGAAACAATGGCAAGA 1239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1920 TAAGGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAAGAGCTGTGATGAAACAATGGCAAGA 1979

Qy  1240 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACTGGATGAAAACAGCCA 1299
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1980 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCA 2039

Qy  1300 AAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTTCGTGTTACAAAGACGTTTGGA 1359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2040 AAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTTCGTGTTACAAAGACGTTTGGA 2099

Qy  1360 TAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAGTCTTCTCAACATTAGGTCCCATTT 1419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2100 TAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAGTCTTCTCAACATTAGGTCCCATTT 2159

Qy  1420 GGAAGCCAGTTCTGACCAGTGGAAGCGGTCTGCACCTTTCTGCGAGGAACTTCTCGTGTG 1479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2160 GGAAGCCAGTTCTGACCAGTGGAAGCGGTCTGCACCTTTCTGCGAGGAACTTCTCGTGTG 2219

Qy  1480 GCTACAGCTGAAAGATGATGAGAATTAAGCCGGCACCCATTGGAGGCGACTTTCCAGC 1539
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2220 GCTACAGCTGAAAGATGATGAGAATTAAGCCGGCACCCATTGGAGGCGACTTTCCAGC 2279

Qy  1540 AGTTCAGAAGCAGCAACGATGTACGAAGATTATCAGAAGATGAAGA 1599
           |||||||||||||||||||||||||| ||
Db  2280 AGTTCAGAAGCAGCAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAAAACTAAAGAACC 2339

Qy  1600 TGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTTCTGACAGAGAATATTCAGAAGATGAAGA 1659
```

```
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 59-60; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;

Query Match             93.7%;  Score 2250;  DB 6;  Length 4966;
Best Local Similarity   94.5%;
Matches 2401;  Conservative    0;  Mismatches    0;  Indels  141;  Gaps  1;

Qy      1 ACATGCATTCAACATCGCCAGATATCATTCAATTAGGCATAGAGAAACTACTGATCCTGAAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1357 ACATGCATTCAACATCGCCAGATATCATTCAATTAGGCATAGAGAAACTACTGATCCTGAAGA 1416

Qy     61 TGTTGATACCACCTTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCTCTTCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1417 TGTTGATACCACCTTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCTCTTCCA 1476

Qy    121 AGTTTTGCCTCAACAAGTGACATTGAGCCATCCAGGAAGTGGAAAATGTTGCCAAGGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 AGTTTTGCCTCAACAAGTGACATTGAGCCATCCAGGAAGTGGAAAATGTTGCCAAGGCC 1536

Qy    181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1537 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 1596

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGGACACAGGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 GATCACGGTCAGTCTAGCACAGGGATATGGACACAGGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1656

Qy    301 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTTCACCACCTTGGCAGTTCATTGGAGCCCATT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTTCACCACCTTGGCAGTTCATTGGAGCCCATT 1716

Qy    361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTGATCATTTGGCAGTTCATTGATGGAGAG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTGATCATTTGGCAGTTCATTGATGGAGAG 1776

Qy    421 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836
```

```
Db   2340 TGTAATCATGATGAGTACTTCTTGAGACTGTACGAATATTTCTGACAGACCTTGGAAGG 2399

Qy   1660 ACTAGAGAAACTTCTACCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGT 1719
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2400 ACTAGAGAAACTCTACCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGT 2459

Qy   1720 CACTCGGCTTCTACGAAAGCAGCGGCTCAATACTGAGTGGGAAAAATTGAACCT 1779
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460 CACTCGGCTTCTACGAAAGCAGCGGCTCAATACTGAGTGGGAAAAATTGAACCT 2519

Qy   1780 GCACTCCGCTGACTGGCAGGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCA 1839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2520 GCACTCCGCTGACTGGCAGGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCA 2579

Qy   1840 AGAGGCCACGATGAGCTGACCTGAAGCTCAAGCTGAGGTGATCAAGCTGGATCCTG 1899
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2580 AGAGGCCACGATGAGCTGGACCTGAAGCTCAAGCTGAGGTGATCAAGCTGGATCCTG 2639

Qy   1900 GCAGCCCGTGGGCGATCTCCTCATTGACTTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 1959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2640 GCAGCCCGTGGGCGATCTCCTCATTGACTTCTCTCCAAGATCACCTCGAGAAAGTCAAGGC 2699

Qy   1960 ACTTCGAGAGAAATTGCGGCCTTGAAAGAAGACGTGAGCCCACGTCAATGACCTTGCTCG 2019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700 ACTTCGAGAGAAATTGCGGCCTTGAAAGAAGACGTGAGCCCACGTCAATGACCTTGCTCG 2759

Qy   2020 CCAGCTTACCACCTTTGGGCATTCAGCTGCTCACCACGTATAACCTCAGCACTCTGGAAGACCT 2079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2760 CCAGCTTACCACCTTTGGGCATTCAGCTGCTCACCACGTATAACCTCAGCACTCTGGAAGACCT 2819

Qy   2080 GAACACCAGATGAGGAAGTTCTGCAGGTGGCCGTGCAGGACCTCAGGCAGCTGCATGA 2139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2820 GAACACCAGATGAGGAAGTTCTGCAGGTGGCCGTGCAGGACCTCAGGCAGCTGCATGA 2879

Qy   2140 AGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCGTCTGTCCAGGGTCC 2199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2880 AGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCGTCTGTCCAGGGTCC 2939

Qy   2200 CTGGGAGAGAGAGCCCATCTGCCCAAACAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 2259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2940 CTGGGAGAGAGAGCCCATCTGCCCAAACAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 2999

Qy   2260 AACTTGCTGGGACCATCCCAAAATGACAGAGAGCTCTTTAGCTGACCTGACCTGAATAA 2319
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3000 AACTTGCTGGGACCATCCCAAAATGACAGAGAGCTCTTTAGCTGACCTGACCTGAATAA 3059

Qy   2320 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAGAAACTCCGAGAGGCCCTTTG 2379
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3060 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAGAAACTCCGAGAGGCCCTTTG 3119

Qy   2380 CTTGGATCTCTTGAGCCTGTCA 2401
          ||||||||||||||||||||||
Db   3120 CTTGGATCTCTTGAGCCTGTCA 3141
```

```
RESULT 6
AAD37256
ID   AAD37256 standard; DNA; 4966 BP.
XX
AC   AAD37256;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
```

```
Qy    481  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAGTGGTGAAAGA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAGTGGTGAAAGA 1896

Qy    541  CCAGTTTCATACTCATGAGGGGTACATGAGCGATTGGATTGACAGCCCATCAGGGGCCGGGTTGG  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897  CCAGTTTCATCTCATGAGGGGTACATGAGCGATTGGATTGACAGCCCATCAGGGGCCGGGTTGG 1956

Qy    601  TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957  TAATATTCTACAATTGGGAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2016

Qy    661  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017  AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

Qy    721  TAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077  TAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 2136

Qy    781  GAAAGAGTTGAATGACTGGCTAACAAGAAGAGAGAAACAAGGAAAAATGGAGGAAGA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137  GAAAGAGTTGAATGACTGGCTAACAAACAGAAGAAACAAGGAAAAATGGAGGAAGA 2196

Qy    841  GCCTCTTGGACCTGATCTTGAAGACAAGTCAGGGTCAATTCTCTCCACTCATGGTCGTGGTAGT  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197  GCCTCTTGGACCTGATCTTGAAGACACAAGTCAGGGTCAATTCTCTCACTCATGGTCGTGGTAGT 2256

Qy    901  AGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTCAAT TCTCTCACTCACTCACAGGTGCTTCA  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257  AGAAGATCTAGAACAAGAACAAGCAAGTCAGGGTCAAT TCTCTCACTCACATAAGGTGCTTCA 2316

Qy    961  TGATGAATCTAGTGGACATCACGCAACTGCTGTCTTGGAAGAACAACTTAAGGTATTCGG  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317  TGATGAATCTAGTGGACATCACGCAACTGCTGTCTTGGAAGAACAACTTAAGGTATTGGG 2376

Qy   1021  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGGCTGGGTTCTTTACAAGAC--  1078
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGGCTGGGTTCTTTTACAAGACCA 2436

Qy   1079  ---------------------------------------------------------  1078
           
Db   2437  GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTTCTCCTACTCAGACTGTTACTCT 2496

Qy   1079  ---------------------------------------------------------  1078
           
Db   2497  GGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAAATGCCATCTTC 2556

Qy   1079  --------------ACTCATAGATTACTGCAACAAGTCGCCATTGGTCCAGACCTGAAAA  1119
                         ||||||||||||||||||||||||||||||||||||||||||||
Db   2557  CTTGATGTTGGAGGTAGCTACTGCAACAAGTCGCCATTGGTTCCCCGGTCCCCTGGACCTGAAAA 2616

Qy   1120  GTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAATGTGCCTACTATATCACCACGATGCTACCCG  1179
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2617  GTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAATGTGCCTACTATATCAACCACGATGCTACCCG 2676

Qy   1180  TAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAACAATGGCAAGA  1239
           |||||||||||||||||||||||||||||||||||||||||||||
Db   2677  TAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAACAATGGCAAGA 2736

Qy   1240  CCTCCAAGGTGAAATTGAAGTCCACCAGATGTTTATCACAACCTGGATGAAAAACAGCCA  1299
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2737  CCTCCAAGGTGAAATTGAAGTCCACCAGATGTTTATCACAACCTGGATGAAAAACAGCCA 2796

Qy   1300  AAAAATCCTGAGATCCTGAAGTTCCGGAAAAAGTCTCTCAACAATCATTAGGTCCCATTT  1359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2797  AAAAATCCTGAGATCCTGAAGTTCCGGAAAAAGTCTCTCAACAATCATTAGGTCCCATTT 2856

Qy   1360  TAACAATGAACTTCAAGTGGAGTGGAAGCGTCTGACCTGTGAACTT  1419
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   2857  TAACAATGAACTTCAAGTGGAGTGGAAGCGTCTGCACCTTCTCTGAGCCTGTG 2916

Qy   1420  GGAAGCCAGTTCTGACCAGTGGAAGCGTCTGACCTTTCTCTGCGGACCTTCTCTGGTGTG 2976
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2917  GGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCGGACCTTCTCTGGTGTG 2976

Qy         GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCCAGCCGGCACCTATTGGAGGCGACTTTCCAGC  1539
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2977  GCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCCGGCAGCCGGCACCTATTGGAGGCGACTTTCCAGC 3036

Qy         AGTTCAGAAGCAGACGACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACC  1599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3037  AGTTCAGAAGCAGACGACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACC 3096

Qy         TGTAATCATGATGACTCTTGACTGTACGAATATTTCTGACAGAGCGCCTTTGGAGG  1659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3097  TGTAATCATGATGACTCTTGACTGTACGAATATTTCTGACAGAGCGCCTTTGGAGG 3156

Qy         ACTAGAGAAACTCTACCAGGAGGCCCAGAGGGTGCCTCCTGAGGAGGAGCCCAGAATGT  1719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3157  ACTAGAGAAACTCTACCAGGAGGCCCAGAGGGTGCCTCCTGAGGAGGAGCCCAGAATGT 3216

Qy         CACTCGGCTTCTACGAAAGCAGCTGTCAATACTGAGTGAGGAAAAATTGAACCT  1779
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3217  CACTCGGCTTCTACGAAAGCAGCTGTCAATACTGAGTGAGGAAAAATTGAACCT 3276

Qy         GCACTCGCTGACTGGCAGAGAAAATGAGACCCTTGAAAGACTCCAGGAACTTCA  1839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3277  GCACTCGCTGACTGGCAGAGAAAATGAGACCCTTGAAAGACTCCAGGAACTTCA 3336

Qy         AGAGAGCCACGGATGAGTCGACCTCAAGCTGACCTGAGCTGATCAAGGATCCTG  1899
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3337  AGAGAGCCACGGATGAGTCGACCTCAAGCTGACCTGAGCTGATCAAGGATCCTG 3396

Qy         GCAGCCGTGTGGGCGTGGGCCGATCCTCCATTGAGAATCTCAGAATCACCTCAAGGC  1959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3397  GCAGCCGTGTGGGCGTGGGCCGATCCTCCATTGAGAATCTCAGAATCACCTCAAGGC 3456

Qy         ACTTCGAGAGGAAATTGCGCCTCTGAAAGAGACGTGAGCCACGTCAATGACCTTGCTCG  2019
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3457  ACTTCGAGAGGAAATTGCGCCTCTGAAAGAGACGTGAGCCACGTCAATGACCTTGCTCG 3516

Qy         CCAGTTACCACTTTGGGCATTCAGCTCTCAGCAGCTCTGAAGACCT  2079
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   3517  CCAGTTACCACTTTGGGCATTCAGCTCTCAGCAGCTCTGAAGACCT 3576

Qy         GAACACCAGATGGAAGCTTCGCAGGTGCGTCAGGTGCGTCGAGCAGCTGCATGA  2139
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3577  GAACACCAGATGGAAGCTTCGCAGGTGCGTCAGGTGCGTCGAGCAGCTGCATGA 3636

Qy         AGCCCACAGGGACTTTGGTCCAGAACACTTCTTCCACGTCGTGTCCAGGGTCC  2199
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3637  AGCCCACAGGGACTTTGGTCCAGAACACTTCTTCCACGTCGTGTCCAGGGTCC 3696

Qy         CTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC  2259
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3697  CTGGGAGAGAGCCATCTCGCCCAAACAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 3756

Qy         AACTTGCTGGGACCATCCAAAATGACAGAGTCTACCAGTCTTTAGCTGACCTGAATAA  2319
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3757  AACTTGCTGGGACCATCCAAAATGACAGAGTCTACCAGTCTTTAGCTGACCTGAATAA 3816

Qy         TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAAGGCCCTTTG  2379
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3817  TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAAGGCCCTTTG 3876

Qy         CTTGGATCTCTTGAGCCTGTCA  2401
           ||||||||||||||||||||||
Db   3877  CTTGGATCTCTTGAGCCTGTCA 3898
```

```
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Cytomegalovirus.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0207777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 67-68; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a
CC   cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ   Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;

Query Match             93.7%;  Score 2250;  DB 6;  Length 4990;
Best Local Similarity   94.5%;  Pred. No. 0;
Matches 2401;  Conservative  0;  Mismatches  0;  Indels 141;  Gaps  1;
```

```
Qy    301 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCACCTGACCCTACACGGAGCCCATT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCACCTGACCCTACACGGAGCCCATT 1740

Qy    361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1800

Qy    421 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1860

Qy    481 TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCTGAGGACACATTGCAAGCACAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1920

Qy    541 CCAGTTTCATATCATGAGGGTACATGATGAATTGACAGCCCATCAGGCGGGTTGG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CCAGTTTCATATCATGAGGGTACATGATGAATTGACAGCCCATCAGGCGGGTTGG 1980

Qy    601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA 2040

Qy    661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2100

Qy    721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGATCAGAAACT  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGATCAGAAACT 2160

Qy    781 GAAAGAGTTGAATGACTGCTACAACAAACAGAGAAGAACAGAAGAAGACAAGGAAGA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GAAAGAGTTGAATGACTGCTACAACAAACAGAGAAGAACAGAAGAAGACAAGGAAGA 2220

Qy    841 GCCTCTTGGACCTCGATCTTGAAGACCTAAAACGCCAAGTACACACAGTAAGGTGCTTCA  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GCCTCTTGGACCTCGATCTTGAAGACCTAAAACGCCAAGTACACACAGTAAGGTGCTTCA 2280

Qy    901 AGAAGATTCTAGAACCAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 AGAAGATTCTAGAACCAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 2340

Qy    961 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTGGAAGACAACTTAAGGTATTGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TGATGAATCTAGTGGAGATCACCGCAACTGCTGCTTTGGAAGACAACTTAAGGTATTGGG 2400

Qy   1021 AGATCGATGGGCAAACATCTGTAGATGACGACAGAAGACCGCTGGGTCTTTACAAGAC-- 1078
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGATCGATGGGCAAACATCTGTAGATGACGACAGAAGACCGCTGGGTCTTTACAAGACCA 2460

Qy   1079 ---------------------------------------------------------- 1078
Db   2461 GCCTGACCTAGCTCCTGACTGACCACTACTTGGAGCCTTCCTCCTACTCAGACTGTTACTCT 2520

Qy   1079 ---------------------------------------------------------- 1078
Db   2521 GGTGACACAACCTGTGGTTACTAAGGAAGCTGCCATCCCAAACTAGAAATGCCATCTTC 2580

Qy   1079 ----------ACTCATAGATTACTGCAACAGTTCCCCTGACCTGAAAA 1119
                    ||||||||||||||||||||||||||||||||||||||
Db   2581 CTTGATGTTGGAGTGACTACCTTACTGCAACAGTTCCCCTGACCTGAAAA 2640

Qy   1120 GTTTCTTGCCTGGCTTACAGAGATGCGAAACAACTGCCAATGCCACTATTCTCAACA 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GTTTCTTGCCTGGCTTACAGAGATGCGAAACAACTGCCAATGCCACTATTCTCAACA 2700

Qy   1180 TAAGGAAAGGCTCCAGAAGAGACTTCCAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGA 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TAAGGAAAGGCTCCAGAAGAGACTTCCAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGA 2760
```

```
Qy      1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 1440

Qy     61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1500

Qy    121 AGTTTGCCTCAACAAGTGAGCATTGAGCAGCCATCCAGGAAGTGCCAAGGCC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AGTTTGCCTCAACAAGTGAGCATTGAGCAGCCATCCAGGAAGTGCCAAGGCC 1560

Qy    181 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCATCAAATGCCACTATTCTCAACA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCATCAAATGCCACTATTCTCAACA 1620

Qy    241 GATCACGGTCAGTCTAGCACAGGGGATATGAGACGGATATGAGAGAAGAACTTCTTCCCCTAAGCCTCGATTCAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GATCACGGTCAGTCTAGCACAGGGGATATGAGACGGATATGAGAGAAGAACTTCTTCCCCTAAGCCTCGATTCAA 1680
```

Db   3841 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCGAAGGCCCTTTG 3900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2380 CTTGGATCTCTTGAGCCTGTCA 2401
           ||||||||||||||||||||||
Db   3901 CTTGGATCTCTTGAGCCTGTCA 3922

RESULT 8
AAD37230
ID   AAD37230 standard; DNA; 4182 BP.
XX
AC   AAD37230;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta4173.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0207777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 43-44; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
CC   hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC   hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 U; 0 Other;

Query Match   86.1%;   Score 2067;   DB 6;   Length 4182;
Best Local Similarity   88.1%;   Pred. No. 0;
Matches 2401   Conservative 0;   Mismatches 0;   Indels 324;   Gaps 1;

Qy   1 ACATGCATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600 ACATGCATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 659

Qy   61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACTACATCACTCTCTTCCA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACTACATCACTCATCTTCCA 719

Qy 1240 CCTCCAAGGTGAAATTGAAGCTCACCACAGATGTTATCAACAACTGGATGAAAACAGCCA 1299
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 CCTCCAAGGTGAAATTGAAGCTCACCACAGATGTTATCACAACAACCTGATGAAAACAGCCA 2820

Qy 1300 AAAAATCCTGGAAGTTCCGATGATGCAGTCCTGTTACAAGACGTTTGGA 1359
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 AAAAATCCTGGAAGTCCCTGGATGATGCAGTCCTGTTACAAGACGTTTGGA 2880

Qy 1360 TAACATGAACTTCAAGTGGGATGGAGTGAACTTCGGAAAAGTCTCTCAACATTAGGTCCCATTT 1419
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 TAACATGAACTTCAAGTGGGATGGAGTGAACTTCGGAAAAGTCTCTCAACATTAGGTCCCATTT 2940

Qy 1420 GGAAGCCAGTTCTGACCAGTGGAAGGGTCTGCACCTTCTGCGAGGAACTTCTGGTGTG 1479
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 GGAAGCCAGTTCTGACCAGTGGAAGGGTCTGCACCTTCTGCGAGGAACTTCTGGTGTG 3000

Qy 1480 GCTACAGCTGAAAGATGAATGAAT TAAGCCGGCACCTATTGGAGGCGACTTTCCAGC 1539
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 GCTACAGCTGAAAGATGAATGAATTAAGCCGGCACCTATTGGAGGCGACTTTCCAGC 3060

Qy 1540 AGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAAAACTAAAGAACC 1599
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 AGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAAAACTAAAGAACC 3120

Qy 1600 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTGGAAGG 1659
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 TGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTGGAAGG 3180

Qy 1660 ACTAGAGAAACTCTACCAGGAGCGCCCAGAGAGCTGCCTCCTGAGGAGCAGAGAGAATGT 1719
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 ACTAGAGAAACTCTACCAGGAGCTGCCTCCTGAGGAGCAGAGAGCCCAGAATGT 3240

Qy 1720 CACTCGGCTTCTACGAAAGCAGCGTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCT 1779
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 CACTCGGCTTCTACGAAAGCAGCGTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCT 3300

Qy 1780 GCACTCCGCTGACTCGCTGCACAGAAAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCA 1839
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 GCACTCCGCTGACTCGCTGCACAGAAAAAATAGATGAGACTCCCTTGAAAGACTTCCAGGAACTTCA 3360

Qy 1840 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGTGAGGTGATCAAGGGATCCTG 1899
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 AGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGTGAGGTGATCAAGGGATCCTG 3420

Qy 1900 GCAGCCCGTGGGCGCATCTCCCATTGACTTCTCCAAGATCACTCGAGAAGTCAAGGC 1959
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GCAGCCCGTGGGCGCCGATCTCCTCCATTGACTCACTCGAGAAGTCAAGGC 3480

Qy 1960 ACTTCGAGGAGAAATTGCGCGCTCTGAAAAGAACCTGAGCCACGTCAATGACCTTGCTCG 2019
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481 ACTTCGAGGAGAAATTGCGCGCTCTGAAAAGAACCTGAGCCACGTCAATGACCTTGCTCG 3540

Qy 2020 CCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGAAGACCT 2079
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 CCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGAAGACCT 3600

Qy 2080 GAACACCAGATGGAAGCTTCGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 2139
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 GAACACCAGATGGAAGCTTCGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGA 3660

Qy 2140 AGCCCACCAGGAGACTTTGGTCCAGCATCTCGAGCACTTCTTCCACGTCTCGTCCAGGGTCC 2199
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 AGCCCACCAGGAGACTTTGGTCCAGCATCTCGAGCACTTCTTCCACGTCTCGTCCAGGGTCC 3720

Qy 2200 CTGGGAGGAGAGCAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAAC 2259
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721 CTGGGAGGAGAGCAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAAC 3780

Qy 2260 AACTTGCTGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 2319
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 AACTTGCTGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA 3840

Qy 2320 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCGAAGGCCCTTTG 2379

```
Qy    121 AGTTTGCCTCAACAAGTGAGCATTGAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 AGTTTGCCTCAACAAGTGAGCATTGAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC  779

Qy    181 ACCTAAAGTGACTAAAGAGAACAGATTTCAGTTACATCACTATTCTCAACA  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 ACCTAAAGTGACTAAAGAGAACAGATTTCAGTTACATCACTATTCTCAACA  839

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA  899

Qy    301 GAGCTATGCCTACACACAGGCTGTCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 GAGCTATGCCTACACACAGGCTGTCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT  959

Qy    361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1019

Qy    421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079

Qy    481 TGCTGAGGACACTTGCAAGCACAAGGACAGATTTCTAATGATGTGGAAGTGGTGAAAGA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACTTGCAAGCACAAGGACAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy    541 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGGTTGG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGGTTGG 1199

Qy    601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATATCAGAAGATGAAGA  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATATCAGAAGATGAAGA 1259

Qy    661 AACTGAAGTACAAGAGCAGAGAATGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAGAGAATGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy    721 TAGCATGAAAGCAATTACAATTACATAGAGTTTTAATGGATCTTCCAGAATCAGAAACT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 TAGCATGAAAGCAATTACAATTACATAGAGTTTTAATGGATCTTCCAGAATCAGAAACT 1379

Qy    781 GAAAGAGTTGAATGACCTTGAAGCAATTACAAAAACGCCAAGTACAACACATAAGGTGCTTCA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTTGAATGACCTTGAAGCAATTACAAAAACGCCAAGTACAACACATAAGGTGCTTCA 1439

Qy    841 GCCTCTTGGACCTGATCTTGAAGACCTAACAGAGAAACGCCAAGTACAAGGAGGAAGA  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 GCCTCTTGGACCTGATCTTGAAGACCTAACAGAGAAACGCCAAGTACAAGGAGGAAGA 1499

Qy    901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGTAGT 1559

Qy    961 TGATGAATCTAGTGGACTCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1560 TGATGAATCTAGTGGACTCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1619

Qy   1021 AGATCGATGGGCAAAACATCTGTAGATGGACAGCAGAGAAAATGGACCGTGGGTTCTTTTACAAGAC 1078
          |||||||||||||||||||||||||||||||||||||||||||||||||||---------
Db   1620 AGATCGATGGGCAAAACATCTGTAGATGGACAGCAGAGAAAATGGACCGTGGGTTCTTTTACAAGACAT 1679

Qy   1079 ------------------------------------------------------- 1078
          
Db   1680 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTTGCCTTTTTAGTGCAATGGCTTTCAGA 1739

Qy   1740 AAAGAAGAATGCAGTGAACAAGATTCACCACAACTGGCTTTAAAGATCAAAAATGAAATGTT 1799
          
Db   1079 ------------------------------------------------------- 1078

Qy   1079 ------------------------------------------------------- 1078
```

```
Db   1800 ATCAAGTCTTCAAAAACTGGCCGTTTAAAGCGGATCTAGAGAAAAAGCAATCCAT 1859
          
Qy   1079 ------------------------------------------------------- 1078

Db   1860 GGGCAAACTGTATTCACTCAACAGATCTCTTTCAACACTGAAGAATAAGTCAGTGAC 1919
          
Qy   1079 ------------------------------------------------------- 1078

Db   1920 CCAGAAGACGGAGCATGGCTCGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 1979
          
Qy   1079 ----------ACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA 1116
          
Db   1980 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA 2039

Qy   1117 AAAGTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 1176
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2040 AAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTAC 2099

Qy   1177 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 1236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2100 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA 2159

Qy   1237 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1296
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2160 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2219

Qy   1297 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTTCCGATGATGCAGTTCCTT 1356
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2220 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTTCCGATGATGCAGTTT 2279

Qy   1357 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTCCAACATTAGGTCCCA 1416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2280 GGATAACATGAACTTCAAGTGAGTGAACTTCGGAAAAAGTCTCCAACATTAGGTCCCA 2339

Qy   1417 TTTGGAAGCCAGTTCTGACCAGTCGTGGAAGCGTCTGACCAGAACTTCTGGT 1476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2340 TTTGGAAGCCAGTTCTGACCAGTCGTGGAAGCGTCTGACCAGAACTTCTGGT 2399

Qy   1477 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGGCACCTATTGGAGGCGACTTTCC 1536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2400 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGGCACCTATTGGAGGCGACTTTCC 2459

Qy   1537 AGCAGTTCAGAGCAGCAGAACGATGTACATAGGGCCTTCAAGAGGGATTGAAAACTAAAGA 1596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460 AGCAGTTCAGAGCAGCAGAACGATGTACATAGGGCCTTCAAGAGGGATTGAAAACTAAAGA 2519

Qy   1597 ACCTGTAATCATGAGTACTCTTGAGACTGTACGACTGCTGCCTCCTGAGGGAGAGCCCAGAA 1656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2520 ACCTGTAATCATGAGTACTCTTGAGACTGTACGACTGCTGCCTCCTGAGGGAGAGCCCAGAA 2579

Qy   1657 AGGACTAGAGAAACTTCACCAGGAGCCCAAGCTGAACTTCGAGCTGAATCGAGAGGGATC 1716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2580 AGGACTAGAGAAACTTCACCAGGAGCCCAAGCTGAACTTCGAGCTGAATCGAGAGGGATC 2639

Qy   1717 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1776
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2640 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 2699

Qy   1777 CCTGCACTCCGCTGACTGGCAGCAGAGAAAAATGATGACCCTTGAAAGACTTCCAGGAACT 1836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700 CCTGCACTCCGCTGACTGGCAGCAGAGAAAAATGATGATGACCCTTGAAAGACTTCCAGGAACT 2759

Qy   1837 TCAAGAGGGCCACGGATGAGCTGGACCTCGGCCAAGCTGAGTCGAGGGATCAAGGGGATC 1896
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2760 TCAAGAGGGCCACGGATGAGCTGGACCTCGGCCAAGCTGAGTCGAGGGATCAAGGGGATC 2819

Qy   1897 CTGGCAGCCCGTGGGCGAATCCTCCTCATTGACTCTTCCAAGATCACCTCCGAAAGTCAA 1956
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2820 CTGGCAGCCCGTGGGCGAATCCTCCTCATTGACTCTTCCAAGATCACCTCCGAAAGTCAA 2879

Qy   1957 GGCACTTCGAGGAGAAATTGCGCCCTCGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2016
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression control
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;

Query Match   86.1%; Score 2067; DB 6; Length 5149;
Best Local Similarity   88.1%; Pred. No. 0;
Matches 2401;   Conservative   0;   Mismatches   0;   Indels 324;   Gaps   1;

Qy   1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   1357 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTACTCGATCCTGAAGA 1416

Qy   61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1417 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1476

Qy   121 AGTTTGCCTCAACAAGTGAGCATTGAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 AGTTTGCCTCAACAAGTGAGCATTGAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   1536

Qy   181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAAATGACTATTCTCAACA   240
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   1537 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGACTATTCTCAACA 1596

Qy   241 GATCACGGTCAGTCTAGCACCAGGGATATGGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 GATCACGGTCAGTCTAGCACCAGGGATATGGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1656

Qy   301 GAGCTATGCCTACACACAGCCTGCTATGTCACCACCTTCACCGGAGCGCCCATT   360
          |||||||||||||| | |||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCCTACCACCAGCCTGCTATGTCACCACCTTCACCGGAGCGCCCATT   1716

Qy   361 TCCTTCACAGCATTTGGAGCTCCTGAGACAAGCAGTCATTGGCAGTTCATTGATGAGAG   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTTGGAGCTCCTGAGACAAGCAGTCATTGGCAGTTCATTGATGAGAG   1776

Qy   421 TGAAGTAAACTGGACCGTTATCATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   1777 TGAAGTAAACTGGACCGTTATCATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

Qy   481 TGCTGAGGACACATTGCAAGCACAAGAGAGATTCTTAATGATGTGGAAGTGGTGAAAGA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGAGAGATTCTTAATGATGTGGAAGTGGTGAAAGA 1896

Qy   541 CCAGTTCATACTCATGAGGGGTACATGATGGCAGCCCATCAGGGGCCGGGTTGG   600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTCATACTCATGAGGGGTACATGATGGCAGCCCATCAGGGGCCGGGTTGG   1956

Qy   601 TAATATTCTACAATTGGGAAGTAAGCTGATTGATTGGAACACAGGAAAAATTATCAGAAGATGAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGATTGGAACACAGGAAAAATTATCAGAAGATGAAGA 2016

Qy   661 AACTGAAGTACAAGAGCTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGGGGAATGCCTCAGGGTAGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGGGGAATGCCTCAGGGTAGC 2076

Qy   721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGCATCCAGAATCAGAAACT   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGCATCCAGAATCAGAAACT   2136

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAGAAAAATGGAGGAGGA   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAGAGAAAAATGGAGGAGGA   2196

Qy   841 GCCTCTTGGACCTGATCTTCGAAGACCTAAAACGCCAAGTACAACATACAACATAAGGTGCTTCA   900

Db   2880 GGCACTTCGGAGGAGAAATTGCGCCTGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2939

Qy   2017 TCGCCAGCTACCACTTGGGCATTCAGCTCTCACCGTATAACCTGACCACTCTGGAAGA 2076
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2940 TCGCCAGCTACCACTTGGGCATTCAGCTCTCACCGTATAACCTGACCACTCTGGAAGA 2999

Qy   2077 CCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCCGTCGAGGACCGTCAGGCAGCTGCA 2136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3000 CCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCCGTCGAGGACCGTCAGGCAGCTGCA 3059

Qy   2137 TGAAGCCACCAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCTGTCCAGGG 2196
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3060 TGAAGCCCACCAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTCTGTCCAGGG 3119

Qy   2197 TCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCA 2256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3120 TCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCA 3179

Qy   2257 AACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTCACCAGTGCTTTAGCTGACCTGAA 2316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3180 AACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTCACCAGTGCTTTAGCTGACCTGAA 3239

Qy   2317 TAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCT 2376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3240 TAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCT 3299

Qy   2377 TTGCTTGGGATCTCTTGAGCCTGTCA 2401
          ||||||||||||||||||||||||||
Db   3300 TTGCTTGGGATCTCTTGAGCCTGTCA 3324

RESULT 9
AAD37255
ID   AAD37255 standard; DNA; 5149 BP.
XX
AC   AAD37255;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 57-59; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

RESULT 10
ABK81999
ID   ABK81999 standard; DNA; 5462 BP.
XX
AC   ABK81999;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaR2-R21+H3.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200229056-A2.
XX

PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Disclosure; Fig 14; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 U; 0 Other;

Query Match   76.6%; Score 1839.4; DB 6; Length 5462;
Best Local Similarity   88.1%; Pred. No.: 0;
Matches 2115; Conservative   0; Mismatches   0; Indels 210; Gaps   2;

Qy   1 ACATGCATTCAACATCGCCAGATATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
           ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   858

Qy   61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCCTTCA   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   859 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCCTTCCA   918

Qy   121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   919 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   978

Qy   181 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   979 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA   1038

Qy   241 GATCACGGGTCAGTCAGCACAGGGATATGAGAGAAACTTCTCCCCTAAGCCTCGATTCAA   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 GATCACGGGTCAGTCAGCACAGGGATATGAGAGAAACTTCTCCCCTAAGCCTCGATTCAA   1098

Qy   301 GAGCTATGCCTACCACACACAGGCTGCTTATGTCACCACCTCGACCCTACACCGGAGCCATT   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACCACACACAGGCTGCTTATGTCACCACCTCGACCCTACACCGGAGCCCATT   1158

Qy   361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCATCATTGGCAGTTCATTGATGGAGAG   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCATCATTGGCAGTTCATTGATGGAGAG   1218

Qy   421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC   1278

Qy   481 TGCTAGGACACATTGCAAGGACACAAGGAGAGATTTCTAATGATGTGGTGAAGTGGTGAAAGA   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTAGGACACATTGCAAGGACACAAGGAGAGATTTCTAATGATGTGGTGAAGTGGTGAAAGA   1338

Qy   541 CCAGTTTCATATCTCATGAGGGGTACATGATGGATTGACAGCGCCATCAGGCGCCGGGTTGG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATATCTCATGAGGGGTACATGATGGATTGACAGCGCCATCAGGCGCCGGGTTGG   1398

Qy   601 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   1458

Qy   661 AACTGAAGTACAGAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAGAGAGCAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC   1518

Qy   721 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   780
           ||||||||||||||||||||||||||||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATGCTCCTGACTG-------------------   1558

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAACAAGGAAAATGGAGGAAGA   840
                                                              --AC   1560
Db   1559 --

Qy   841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA   900
           |||||||||||||| ||||| | || || || || |||||||||||||||
Db   1561 CACTATTGGAGCCTCTTCCTACTCAGACTGTTACTCTGGTGACACAACCGTGGTTACTAA   1620

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACATCACATGGTGGTGTAGT   960
           ||||||| |||||||||||||||||||||||||||||||||||||--------
Db   1621 GGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTCCTTGATGTTGGAG-------   1669

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTGCGTTGCTTTGGAGAACAACTTAAGGATTGGG   1020

Db   1670 --   1669

Qy   1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCGTGGGTTCTTTTACAGACAC   1080

Db   1670 --   1669

Qy   1081 TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGA   1140
           | ||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1670 -CATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGA   1728

Qy   1141 AGTGAAACAACTGCCAATGTCCTACAGGATGCCAAGGTCTACCCGTAAGGAAAGGTCTCCTAGAAGA   1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1729 AGTGAAACAACTGCCAATGTCCTACAGGATGCCAAGGTCTACCCGTAAGGAAAGGTCTCCTAGAAGA   1788

Qy   1201 CTCCAAGGGAGTAAAGACTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC   1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1789 CTCCAAGGGAGTAAAGACTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC   1848

Qy   1261 TCACACAGATGTTTATCACAACCTGATGAAAAATCCTGATAAAAATCCTGAGATCCCTGGA   1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1849 TCACACAGATGTTTATCACAACCTGATGAAAAATCCTGATAAAAATCCTGAGATCCCTGGA   1908

Qy   1321 AGTTCCGATGATGCAGTCAGTTGGATAACATCAAAGACGTTGGATAACATGAACATTCAAGTGGAG   1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1909 AGTTCCGATGATGCAGTCAGTTCCTGTTACAAAGACGTTGGATAACATGAACATTCAAGTGGAG   1968

Qy   1381 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTCTGACCAGTG   1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1969 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTCTGACCAGTG   2028

Qy   1441 GAGGCGTTCGCACCTTTCTCTGCAGGAACTTCGGTGTGGGTACAGCTGAAAGATGATGA   1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2029 GAAGCGTCGCACCTTTCTCTGCGGAACTTCGGTGTGGGTACAGCTGAAAGATGATGA   2088

Qy   1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT   1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2089 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT   2148

Qy   1561 ACATTAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCCTTGA   1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2149 ACATTAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCCTTGA   2208

Qy   1621 GACTGTACGAATATTTCTGACAGAGCACGCCTTTGACAGCAGCAGCGCCTAGAGGACTAGAGAAACTTACCAGGA   1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2209 GACTGTACGAATATTTCTGACAGAGCACGCCTTTGACAGCAGCAGCGCCTAGAGGACTAGAGAAACTTACCAGGA   2268

```
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 50-51; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
CC   hinge H1 and rod R1), 8055-10227 (rods R22, R23 and R24, hinge H4 and CR
CC   domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;

Query Match             72.3%;  Score 1737;  DB 6;  Length 3531;
Best Local Similarity   86.4%;  Pred. No. 0;
Matches 2074;  Conservative  0;  Mismatches  0;  Indels 327;  Gaps  1;

Qy         1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATGCATAGAGAAACTACTCGATCCTGAAGA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATGCATAGAGAAACTACTCGATCCTGAAGA 659

Qy        61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTACTCTTCCA 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       660 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTACTCTTCCA 719

Qy       121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAAATGTTGCCAAGGCC 180
             |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       720 AGTTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAAATGTTGCCAAGGCC 779

Qy       181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAAATGACTATTCTCAACA 240
             |||||||||||||||||||||||||||||||||||||| |||||||| |||||||||
Db       780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGACTACTATTCTCAACA 839

Qy       241 GATCACGGTCAGTCTAGCACAGGGATATGAGACACAGGAACTTCTTCCCCTAAGCCTCGATTCAA 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       840 GATCACGGTCAGTCTAGCACAGGGATATGAGACACAGGAACTTCTTCCCCTAAGCCTCGATTCAA 899

Qy       301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTTACACCACCTTATAG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTTACACCACCTTATT 959

Qy       361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCATTCATTGGCCAGTCATTGATGGAGAG 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCATTCATTGGCCAGTCATTGATGATGGAGAG 1019

Qy       421 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1020 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATATCGTGGCTTCTTTC 1079

Qy       481 TGCTGAGGACACATTGCAAGCACAAGGACAAGGAGCACAAGGACGAGAATTCTAATGATGTGGTGAAAGA 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1080 TGCTGAGGACACATTGCAAGCACAAGGACAAGGAGCACAAGGAGAGAATTTCTAATGATGTGGTGAAAGA 1139
```

```
Qy      1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCACGAAAGCA 1740
             ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      2269 GCCCAGAGAGCTGCCTCCTGAGGAGAGCAGAGCCCAGAATGTCACTCGGCTTCACGAAAGCA 2328

Qy      1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2329 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2388

Qy      1801 AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGATGAGCTGGA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2389 AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGGCCACGGATGAGCTGGA 2448

Qy      1861 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2449 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 2508

Qy      1921 CATTGACTCTCCAGATCACCTCAGGAAGTCAAGGCACTTGCGCC 1980
             ||||||||||| ||||||||| |||||||||||||||||||||||
Db      2509 CATTGACTCTCTCAAGATCACCTCGAGAAGTCAAGGCACTTGCGCC 2568

Qy      1981 TCTGAAAGGAACGTGAGCGACCTCAATGACCTTGCTCGCCAGCTACCACTTTGGGCAT 2040
             |||||| ||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      2569 TCTGAAAGAGAACGTGAGCGACCTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2628

Qy      2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGGAAGCTTCT 2100
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      2629 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2688

Qy      2101 GCAGGTGGCCGTCGAGGACCGAGCTGCAGGCAGTGCATGAAGCCCACAGGGACTTTGGTCC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2689 GCAGGTGGCCGTCGAGGACCGAGCTGCAGGCAGTGCATGAAGCCCACAGGGACTTTGGTCC 2748

Qy      2161 AGCATCTCGAGCACTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGGAGAGGCCATCTCGCC 2220
             ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2749 AGCATCTCAGCACTTCTTTCCACGTCGTCCAGGGTCCCTGGGAGGAGAGGCCATCTCGCC 2808

Qy      2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA 2280
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2809 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA 2868

Qy      2281 AATGACAGAGCTCTTCACCAGTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2869 AATGACAGAGCTCTTCACCAGTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2928

Qy      2341 GACTGCCATGAAAACTCCGAAGACTGCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTC 2400
             |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      2929 GACTGCCATGAAAACTCCGAAGACTCCGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTC 2988

Qy      2401 A 2401
             |
Db      2989 A 2989
```

```
RESULT 11
AAD37238
ID   AAD37238 standard; DNA; 3531 BP.
XX
AC   AAD37238;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Human dystrophin minigene delta3531.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
```

Qy 1621 GACTGTACGAATATTTCTGACAGCAGCAGCCTTTGGAAGGACAGAGAGAAACTCTACCAGGA 1680
```
          ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
```
Db 1893 GACTGTACGAATATTTCTGACAGCAGCAGCCTTTGGAAGGACAGAGAGAAACTCTACCAGGA 1952

Qy 1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1740
```
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
```
Db 1953 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 2012

Qy 1741 GGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCCAGAG 1800
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2013 GGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCCAGAG 2072

Qy 1801 AAAAATAGATGAGACCCTTGAAAGACTTCAGGAGCCACGGATGAGCTGGA 1860
```
          |||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2073 AAAAATAGATGAGACCCTTGAAAGACTTCAGGAGCCACGGATGAGCTGGA 2132

Qy 1861 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGAGTCCTGGCAGCCCGTGGGCAGCCATCCT 1920
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2133 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGAGTCCTGGCAGCCCGTGGGCAGCCATCCT 2192

Qy 1921 CATTGACTCCTCCAAGATCACCTCGAGAAAGTCAATGACCTTGCTCGCCAGCTTACCACTTGGGCAT 1980
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2193 CATTGACTCCTCCAAGATCACCTCGAGAAAGTCAATGACCTTGCTCGCCAGCTTACCACTTGGGCAT 2252

Qy 1981 TCTGAAAGAGAACGTGAGCCACTCTCAGCACTCAGCCAGAGACCTGAACACCAGATGGAAGCTTCT 2040
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2253 TCTGAAAGAGAACGTGAGCCACTCTCAGCACTCAGCCAGAGACCTGAACACCAGATGGAAGCTTCT 2312

Qy 2041 TCAGTCTCACGTATAACCTCAGCAGTCGAGGACCGAGGTCAGGCGAGCTGCATGAGAAGCCACCAGAGGGACTTTGGTCC 2100
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2313 TCAGTCTCACGTATAACCTCAGCAGTCGAGGACCGAGGTCAGGCGAGCTGCATGAGAAGCCACCAGAGGGACTTTGGTCC 2372

Qy 2101 GCAGGTGGCCGTGCGAGGACCGAGGTCAGGCGAGCTGCATGAGAAGCCCAGAGGGACTTTGGTCC 2160
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2373 GCAGGTGGCCGTGCGAGGACCGAGGTCAGGCGAGCTGCATGAGAAGCCCAGAGGGACTTTGGTCC 2432

Qy 2161 AGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGGAGAGCCATCTCGCC 2220
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2433 AGCATCTCAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGGAGAGCCATCTCCCC 2492

Qy 2221 AAACAAAGTGCCCACTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA 2280
```
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
```
Db 2493 AAACAAAGTGCCCACTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA 2552

Qy 2281 AATGACAGAGCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2553 AATGACAGAGCTCACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2612

Qy 2341 GACTGCCATGAAACTCCGAGAACTGCGAGAAGGGCCCTTTGCTTGGATCCTTTGCTTGGATCCTGTGTC 2400
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 2613 GACTGCCATGAAACTCCGAGAACTGCGAGAAGGGCCCTTTGCTTGGATCCTTTGCTTGGATCCTGTGTC 2672

Qy 2401 A 2401

Db 2673 A 2673

RESULT 12
AAD37258
ID   AAD37258 standard; DNA; 4498 BP.
XX
AC   AAD37258;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-3531.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.

Qy 541 CCAGTTCATACTCATGAGGGGTACATGATGGATTTGACAGCGCCATCAGGGCCGGGTTGG 600
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1140 CCAGTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGGTTGG 1199

Qy 601 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1200 TAATATTCTACAAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259

Qy 661 AACTGAAGTACAAGAGCAGATGAATCTCCCAAATTCAGAGATGGGAATGGCTCAGGGTAGC 720
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGAGATGGGAATGGCTCAGGGTAGC 1319

Qy 721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCGAATCAGAAACT 780
```
          |||||||||||||||||||||||||||||||||||
```
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGA-------------------------- 1350

Qy 781 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAACAAGGAGAAAATGAGGAGGAAGA 840
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1350 -- 1350

Qy 841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 900
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 900 -- 900

Qy 901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 960
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1350 -- 1350

Qy 961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1020
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1350 -- 1350

Qy 1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||    ||
```
Db 1351 --AC 1352

Qy 1081 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1140
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1353 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1412

Qy 1141 AGCTGAAACAACTGCCAATGTCCTACCAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1200
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1413 AGCTGAAACAACTGCCAATGTCCAATGTCCTACCCGTAAGGAAAGGCTCCTAGAAGA 1472

Qy 1201 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1473 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1532

Qy 1261 TCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGA 1320
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1533 TCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGA 1592

Qy 1321 AGGTTCCGATGATGGAGTCTCTCAACATTAGGTCCCATTTGGAAGCAGCCAGTTCTGACCAGTG 1380
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1653 AGGTTCCGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCAGCCAGTTCTGACCAGTG 1712

Qy 1381 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCAGCCAGTTCTGACCAGTG 1440
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1712 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCAGCCAGTTCTGACCAGTG 1772

Qy 1441 GAAGCCGTCTGCACCTTTCTCTGCAGGAACTTCTCGTGTGTGCTACAGCTGAAAGATGATGA 1500
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1713 GAAGCCGTCTGCACCTTTCTCTGCAGGAACTTCTCGTGTGTGCTACAGCTGAAAGATGATGA 1772

Qy 1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 1560
```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1773 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 1832

Qy 1561 ACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1833 ACATAGGGCCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1892

```
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0207777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
PR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 62-63; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 U; 0 Other;

Query Match          72.3%;  Score 1737;  DB 6;  Length 4498;
Best Local Similarity 86.4%;  Pred. No. 0;
Matches 2074;  Conservative 0;  Mismatches 0;  Indels 327;  Gaps 1;
```

```
Qy    481 TGCTGAGGACACATTGCAAGCACAAGCAGACAAGGAGAGAGATTTCTAATGATGATGTGAAGTGGTGAAAGA 540
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACAATTGCAAGCACAAGCAGACAAGGAGAGAGATTTCTAATGATGATGTGAAGTGGTGAAAGA 1896

Qy    541 CCAGTTCATACTCCATGAGGGGTACATCATGATGATTTGACAGCCCATCAGGGGCCGGGTTGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTCATACTCCATGAGGGGTACATCATGATGATTTGACAGCCCATCAGGGGCCGGGTTGG 1956

Qy    601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

Qy    661 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGAGTGGAATGCCTCAGGGTAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGAGTGGAATGCCTCAGGGTAGC 2076

Qy    721 TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATGATCTCCAGAATCAGAAACT 780
          |||||||||||||||||||||||||||||
Db   2077 TAGCATGGAAAAACAAAGCAATTACATAGA--------------------------- 2107

Qy    781 GAAAGAGTTGAATGACTGCTAACAAAAACAGAAGAAGAACAAGGAAAATGGAGGAAGA 840
Db   2108 ------------------------------------------------------- 2107

Qy    841 GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 900
Db   2108 ------------------------------------------------------- 2107

Qy    901 AGAAGATTTAGAACAAGAACAAGTCAGGTCAATTCTCCTACTCACATGGTGGTGGTAGT 960
Db   2108 ------------------------------------------------------- 2107

Qy    961 TGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
Db   2108 ------------------------------------------------------- 2107

Qy   1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
          ||
Db   2108 ---------------------------------------------------------AC 2109
```

```
Qy      1 ACATGCATTCAACATCGCCAGATATCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1357 ACATGCATTCAACATCGCCAGATATCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 1416

Qy     61 TGTTGATACCACCTTATCCAGATAAGAAGTCCATCTTAATGTACATCACACATCACTCTTCCA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1417 TGTTGATACCACCTTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCACTCCA 1476

Qy    121 AGTTTTGCCTCAACAAGTGACTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTGCCAAGGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 AGTTTTGCCTCAACAAGTGACTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTGCCAAGGCC 1536

Qy    181 ACCTAAAGTGACTAAAGAGAAGAACATTTTCAGTTACAACATTCAGTTACAACATCAAATGCACTATTCTCAACA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1537 ACCTAAAGTGACTAAAGAGAAGAACATTTTCAGTTACAACATCAAATGCACTATTCTCAACA 1596

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGCACAGGGATATGCACAGGATACACTTCTCCCCTAAGCCTGATTCAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 GATCACGGTCAGTCTAGCACAGGGATATGCACAGGATACACTTCTCCCCTAAGCCTGATTCAA 1656

Qy    301 GAGCTATGCTACCACCACAGGCTCGCTGCTTATGTCACCACCTCACGGCCTACACGGAGCCCATT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCTACACACCACAGGCTCGCTGCTTATGTCACCACCTCACGGCCTACACGGAGCCCATT 1716

Qy    361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCATTTGGCAGTTCATTGATGAGAG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCAAGTCATTGGCAGTTCATTGATGAGAG 1776

Qy    421 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTTATCGTTGTGGCTTCTTTC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTTATCGTTGTGGCTTCTTTC 1836
```

```
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200229056-A2.
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031126.
XX
PR   06-OCT-2000; 2000US-0238848P.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Chamberlain JS, Harper SQ;
XX
DR   WPI; 2002-435334/46.
XX
PT   A composition for preparing therapeutic drugs, has a mini-dystrophin
PT   peptide comprising a specific number of spectrin-like repeat domains, or
PT   a nucleic acid sequence encoding the mini-dystrophin peptide.
XX
PS   Example 6; Fig 13; 145pp; English.
XX
CC   The invention describes a composition comprising a mini-dystrophin
CC   peptide comprising a spectrin-like repeat domain, where the domain
CC   comprises n spectrin-like repeats, and contains no more than n spectrin-
CC   like repeats, where n is an even number between 4-24, or a nucleic acid
CC   encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the
CC   polynucleotide encoding it is useful as a medicament, for preparing a
CC   drug for therapeutic application and in the preparation of a composition
CC   for treatment of muscle disease, e.g. Duchenne's muscular dystrophy
CC   (DMD). This sequence represents a mini-dystrophin sequence of the
CC   invention
XX
SQ   Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 U; 0 Other;

Query Match             71.8%;  Score 1725;  DB 6;  Length 5339;
Best Local Similarity   86.1%;  Pred. No. 0;
Matches 2068;  Conservative   0;  Mismatches   0;  Indels 333;  Gaps   1;

Qy        1 ACATGCATTCAACATGCCAGATATCATTAGGCATAGAAACTACTCGATCCTGAAGA 60
            ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db      799 ACATGCATTCAACATCCGCCAGATATCATTAGGCATAGAGAACTACTCGATCCTGAAGA 858

Qy       61 TGTTGATACCACCTATCCAGATAAGAGTCCATCTTAATGCATCACATCACATCCTTCCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      859 TGTTGATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACATCACATCCTTCCA 918

Qy      121 AGTTTGCTCAACAAGTGACATTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      919 AGTTTGCTCAACAAGTGACATTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 978

Qy      181 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      979 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA 1038

Qy      241 GATCACGGTCAGTCTAGCACCAGGGATATGAGGGATATGAGAACTTCTTCCCCTACCCTA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1039 GATCACGGTCAGTCTAGCACCAGGGATATGAGGGATATGAGAACTTCTTCCCCTACCCTA 1098

Qy      301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACCCTACACGGAGCCATT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACCCTACACGGAGCCATT 1158

Qy      361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGGATGGAGAG 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGGATGGAGAG 1218

Qy      421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTTGGCTTCTTTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTTGGCTTCTTTC 1278
```

Left column alignments:

```
Qy     1561 ACATAGGGCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2590 ACATAGGGCCTTCAAGAGGGAATTGAAACTAAAGAACCTGTTAATCATGAGTACTCTTGA 2649

Qy     1621 GACTGTACGAATATTTCTGACAGAGGACTAGAGAGAGACCTTGAAGAGGACTAGAGAAACTCTACCAGGA 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2650 GACTGTACGAATATTTCTGACAGAGGACTAGAGAGAGACCTTGAAGAGGACTAGAGAAACTCTACCAGGA 2709

Qy     1681 GCCCAGAGAGCAGCTGCCTCCTGAGAGAGAGCCCAGAATGTCACTCGGCTTCTTCACGAAAGCA 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2710 GCCCAGAGAGCAGCTGCCTCCTGAGAGAGAGCCCAGAATGTCACTCGGCTTCTTCACGAAAGCA 2769

Qy     1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2770 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2829

Qy     1801 AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGACTCAAGAGGCCACGGATGAGCTGGA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2830 AAAAATAGATGAGACCCTTGAAAGACTTCAAGAGACTCAAGAGGCCACGGATGAGCTGGA 2889

Qy     1861 CCTCAAGCTGCGCCAAGTGAGGTGATCAAGGGATCCTCGGCGAGCCCGTGGGCGATCTCCT 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2890 CCTCAAGCTGCGCCAAGTGAGGTGATCAAGGGATCCTCGGCGAGCCCGTGGGCGATCTCCT 2949

Qy     1921 CATTGACTCTCCAAGATCACCTCCAGAGAAGTCAAGGCACTTCAGGGAGAAATTGCGCC 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2950 CATTGACTCTCCAAGATCACCTCCAGAGAAGTCAAGGCACTTCAGGGAGAAATTGCGCC 3009

Qy     1981 TCTGAAAGAAGCTGAGCCACGTGCTTCGCCGAGCTTACCACTTGGGCAT 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db     3010 TCTGAAAGAAGCTGAGCCACGTGCTTCGCCGAGCTTACCACTTGGGCAT 3069

Qy     2041 TCAGCTCTTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGGATGGAAGCTTCT 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3070 TCAGCTCTTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGGATGGAAGCTTCT 3129

Qy     2101 GCAGGTGGCCGTCGAGGACCGAGCAGCTCAGGGACTTTGGTCC 2160
            ||||||||||||||||||||||||||||||||||||||||||
Db     3130 GCAGGTGGCCGTCGAGGACCGAGCAGCTCAGGGACTTTGGTCC 3189

Qy     2161 AGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3190 AGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3249

Qy     2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3250 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA 3309

Qy     2281 AATGACAGAGCAGCTTCACCAGGTCTTTAGCTGACCTGAATAATGCTTCAGCTTATAG 2340
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3310 AATGACAGAGCAGCTTCACCAGGTCTTTAGCTGACCTGAATAATGCTTCAGCTTATAG 3369

Qy     2341 GACTGCCATGAAACTCCGAAGACTGCGAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3370 GACTGCCATGAAACTCCGAAGACTGCGAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 3429

Qy     2401 A 2401
            |
Db     3430 A 3430
```

```
RESULT 13
ABK81998
ID   ABK81998 standard; DNA; 5339 BP.
XX
AC   ABK81998;
XX
DT   13-AUG-2002 (first entry)
XX
DE   DNA encoding mini-dystrophin protein deltaR2-R21.
XX
KW   Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW   Duchenne's muscular dystrophy; DMD; dystrophin; ds.
```

```
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 51-52; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or H1 and H4
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to a recombinant control
CC   element. The dystrophin minigene operably linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
CC   hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CC   CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;
```

```
Query Match           70.6%;  Score 1695;  DB 6;  Length 3510;
Best Local Similarity 85.5%;  Pred. No. 0;
Matches 2053;  Conservative 0;  Mismatches 0;  Indels 348;  Gaps 1;
```

```
Qy      1 ACATGCATTCAACATGCCAGATATCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 ACATGCATTCAACATGCCAGATATCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA  659

Qy     61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACATCCTTCCA  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACATCCTTCCA  719

Qy    121 AGTTTTGCCTCAACAAGTGAGCAATTGAAGCCATCCAGGAAGTGAAAATGTTGCCAAGGCC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 AGTTTTGCCTCAACAAGTGAGCAATTGAAGCCATCCAGGAAGTGAAAATGTTGCCAAGGCC  779

Qy    181 ACCTAAAGTGACTAAAGAAGAAAACATTTTCAGTTACCATCATCAAATGCCACTATTCTCAACA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 ACCTAAAGTGACTAAAGAAGAAAACATTTTCAGTTACCATCATCAAATGCCACTATTCTCAACA  839

Qy    241 GATCACGGTCAGTCTAGCACAGGCGTCAGCCGGATATGAGAAGACTTCCTAAGCCTCGATTCAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 GATCACGGTCAGTCTAGCACAGGCGTCAGCCGGATATGAGAAGACTTCTTCCCCTAAGCCTCGATTCAA  899

Qy    301 GAGCTATGCCTACACACAGGCTGCTATGTCACACCTACCCTACGGAGCCCATT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 GAGCTATGCCTACACACAGGCTGCTATGTCACACCTACCCTACGGAGCCCATT  959

Qy    361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGCAAGTCATTTGGCAGTTCATTGATGGAGAG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420

Qy    421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTC  480
```

```
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 63-65; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigene, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;

Query Match             70.6%;  Score 1695;  DB 6;  Length 4476;
Best Local Similarity   85.5%;  Pred. No. 0;
Matches  2053;  Conservative  0;  Mismatches  0;  Indels  348;  Gaps  1;

Qy          1 ACATGCAATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAACTACTCGATCCTCGAAGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1356 ACATGCAATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAACTACTCGATCCTCGAAGA 1415

Qy         61 TGTTGATACCACCTTATCCAGATAAGAAGTCATCTTAATGTACATCACATCACATCACTCCTTCCA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1416 TGTTGATACCACCTTATCCAGATAAGAAGTCATCTTAATGTACATCACATCACATCACTCCTTCCA 1475

Qy        121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1476 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1535

Qy        181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAAATGCAATCATCAAATGCACTATTCTCAACA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1536 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCAAATGCAATCATCAAATGCACTATTCTCAACA 1595

Qy        241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1596 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1655

Qy        301 GAGCTATGCCTACCACAGGCTGCTTATGTCACCAGTTCACCACCTTCACCACCGGAGCCCCATT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1656 GAGCTATGCCTACCACAGGCTGCTTATGTCACCAGTTCACCACCTTCACCACCGGAGCCCCATT 1715

Qy        361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGCAAGTCATTTGGCAGTTCATTGATGGAGAG 420
```

```
Db       1752 ATTAAGCGCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 1811
Qy       1561 ACATAGGGGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1812 ACATAGGGGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1871

Qy       1621 GACTGTACGAATATTTCTGACAGAGCAGCAGCCTTGGAAGGACTAGAGAAACTCTACCAGGA 1680
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1872 GACTGTACGAATATTTCTGACAGAGCAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 1931

Qy       1681 GCCCAGAGAGCTGCCTCCTCGAGGAGAGAGCCCAGAAATGCACTCGGCTTCTACGAAAGCA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1932 GCCCAGAGAGCTGCCTCCTCGAGGAGAGAGCCCAGAAATGCACTCGGCTTCTACGAAAGCA 1991

Qy       1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCCAGAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1992 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTCGCACTCCGCTGACTGGCCAGAG 2051

Qy       1801 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGAGGCCACGGATGAGCTGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2052 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGAGGCCACGGATGAGCTGGA 2111

Qy       1861 CCTCAAGTGCGCCAAGCTGAGTGATCAAGGATCCTGGCACCGCCGTGGGCGATCTCCT 1920
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2112 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGATCCTGGCAGCCCGTGGGCGATCCCCT 2171

Qy       1921 CATTGACTTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTGAGGAGGAGAATTGCGCC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2172 CATTGACTTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTGAGGAGGAGAATTGCGCC 2231

Qy       1981 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTGGGCAT 2040
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Th       2232 TCTGAAAGAGAACGTGAGCGCACGTCAATGACCTTGCTCGCCAGCTTACCACTTGGGCAT 2291

Qy       2041 TCAGCTCCTCACCGTATAACCTCAGCACTCTCAGCACTCTGAAGACACCAGATGAAGCTTCT 2100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2292 TCAGCTCCTCACCGTATAACCTCAGCACTCTCAGCACTCTGAAGACACCAGATGAAGCTTCT 2351

Qy       2101 GCAGGTGGCCGTGGAGGACCGAGTCAGGGACGCTGCAGCGCGAGTCAAGGACTTTGGTCC 2160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2352 GCAGGTGGCCGTGGAGGACCGAGTCAGGGACGCTGCAGCGCGAGTCAAGGACTTTGGTCC 2411

Qy       2161 AGGATCTCAGCACTTCTTTCCACGTCTGTCTGTCCAGGGGTCCCTGGAGAGCAGCCATCCGCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2412 AGGATCTCAGCACTTCTTTCCACGTCTGTCTGTCCAGGGGTCCCTGGAGAGCAGCCATCCGCC 2471

Qy       2221 AAACAAAGTGCCCTACTATATCAACCACCGAGACTCAAACAACTGCTGGGACCATCCCAA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2472 AAACAAAGTGCCCTACTATATCAACCACCGAGACTCAAACAACTGCTGGGACCATCCCAA 2531

Qy       2281 AATGACAGAGCTCACCAGTTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2532 AATGACAGAGCTCACCAGTTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2591

Qy       2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2592 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2651

Qy       2401 A 2401
              |
Db       2652 A 2652
```

```
RESULT 15
AAD37259
ID   AAD37259 standard; DNA; 4476 BP.
XX
XC   AAD37259;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-3510.
```

```
Db  1716 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1775
         TCCTTCAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGATGGAG

Qy   421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC  480
         |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  1776 TGAAGTAAACCTGGACCGTTATCAAACAGCAGTTTAGAGAGTATTATCGTGGCTTCTTTC 1835

Qy   481 TGCTGAGGACACCATTGCAAGCACAACGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1836 TGCTGAGGACACCATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1895

Qy   541 CCAGTTTCATACTCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCCGGGTGG  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1896 CCAGTTTCATCATCATGAGGGGTACATGATGATTGACAGCACGAAAAATTATCAGAAGATGAAGA 1955

Qy   601 TAATATTCTACAATTGGGAGTAAGCTGATTGGACAGGAAAATTATCAGAAGATGAAGA  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1956 TAATATTCTACAATTGGGAGTAAGCTGATTGGACAGGAAAATTATCAGAAGATGAAGA 2015

Qy   661 AACTGAAGTACAGAGCAGATGAATCCTAAATTCAAGATGGAATGCCTCAGGGTAGC  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2016 AACTGAAGTACAGAGCAGATGAATCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 2075

Qy   721 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2076 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTTAATGGATCTCCCAGAATCAGAAACT 2135

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAGAAAATGAGGAGGA  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2136 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAACAAGGAGAAAATGAGGAGGA 2195

Qy   841 GCCTCTTGGACCTGATCTTGAAGACTACAACGCCAAGTACAACACATAAGGTGCTTCA  900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2196 GCCTCTTGGACCTGATCTTGAAGACTACAACGCCAAGTACAACACATAAGGTGCTTCA 2255

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCCACATGGTGGTGGTAGT  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2256 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCCACATGGTGGTGGTAGT 2315

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAGAACAACTTAAGGTATTGGG 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2316 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAGAACAACTTAAGGTATTGGG 2375

Qy  1021 AGAATCGATAGGGCAAACATCTGTAGATGGACAGAAGAGACCGCTGGTTCTTTTACAAGACAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||--
Db  2376 AGATCGATAGGGCAAACATCTGTAGATGGACAGAAGAGACCGCTGGGTTCTTTTACAAGA--- 2432

Qy  1081 TCATAGATTACTGCAACAGTTCCCCCTGGAACTTTTCTTGCCTGGCTTACAGA 1140
         ------------------------------------------------------
Db  2433 ------------------------------------------------------

Qy  1141 AGCTGAAACAACTGCCAATGTCTGCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGA 1200
         ------------------------------------------------------------
Db  2433 ------------------------------------------------------------

Qy  1201 CTCCAAGGGAGTAAAAGAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
         ------------------------------------------------------------
Db  2433 ------------------------------------------------------------

Qy  1261 TCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAAATCCTGAGATCCCTGA 1320
         -----------------------------------------------------------
Db  2433 -----------------------------------------------------------

Qy  1321 AGGTTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380
         -----------------------------------------------------------
Db  2433 -----------------------------------------------------------

Qy  1381 TGAACTTCGGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1440
         -----------------------------------------------------------
Db  2433 -------------------------------------------CAGTTCTGACCAGTG 2447

Qy  1441 GAAGCGGCCTGCACCTTCTCTGCAGGAACTTCTCGTGTGGCTACAGCTGAAAGATGATGA 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  2448 GAAGCGTCTGCACCTTCTCTGCAGGAACTTCTCGTGTGGCTACAGCTGAAAGATGATGA 2507

Qy  1501 ATTAAGCGCAGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAGCAGAACGATGT 1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2508 ATTAAGCGCAGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAGCAGAACGATGT 2567

Qy  1561 ACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCCTGA 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2568 ACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCCTGA 2627

Qy  1621 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGAGACTAGAGAACTCTACCAGGA 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2628 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGAGACTAGAGAACTCTACCAGGA 2687

Qy  1681 GCCCAGAGAGCTGCCTCCTGAGGAGGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2688 GCCCAGAGAGCTGCCTCCTGAGGAGGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 2747

Qy  1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAAACTGCACTCCGCTGACTGGCAGAG 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2748 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAAACTGCACTCCGCTGACTGGCAGAG 2807

Qy  1801 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2808 AAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2867

Qy  1861 CCTCAAGCTGCGCGCCAAGCTGGATGATCCTCGAGCTGATCAAGGCTGTGGGCGATCTCCT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2868 CCTCAAGCTGCGCGCCAAGCTGGATGATCCTCGAGCTGATCAAGGGCCGTGGGCGATCTCT 2927

Qy  1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2928 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 2987

Qy  1981 TCTGAAAGAACGTGAGCCAGTCAATGACCTTCGCCAGCTTACCCACTTTGGGCAT 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2988 TCTGAAAGAACGTGAGCCAGTCAATGACCTTCGCCAGCTTACCCACTTTGGGCAT 3047

Qy  2041 TCAGCTTCTCACCGTATAACCTCAGCACTCTGGAAGACTCTGGAACACCAGATGGAAGCTTCT 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3048 TCAGCTTCTCACCGTATAACCTCAGCACTCTGGAAGACTCTGGAACACCAGATGGAAGCTTCT 3107

Qy  2101 GCAGGTGCCGTGCGAGGACCGAGTCAGGCGAGCGCTGCATGAGAAGCCCACAGGGACTTTGGTCC 2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3108 GCAGGTGGCCGTGCGAGGACCGAGTCAGGCGAGCGCTGCATGAGAAGCCCACAGGGACTTTGGTCC 3167

Qy  2161 AGCATCTCAGCACTTTCTTTCCTTCCTCGTCTCCAGGTCCTCGGGAGAGCCATCTCCGCC 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3168 AGCATCTCAGCACTTTCTTTCCTTCCTCGTCTCCAGGTCCTCGGGAGAGCCATCTCCGCC 3227

Qy  2221 AAACAAAGTGCCCTACTATATCAAACAACTTGCTGGGACCATCCCAA 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3228 AAACAAAGTGCCCTACTATATCAAACAACTTGCTGGGACCATCCCAA 3287

Qy  2281 AATGACAGAGCCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3288 AATGACAGAGCCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3347

Qy  2341 GACTGCCATGAAAACTCCGAGAACTCCGAGAGATGCGAGAGAGGCCCTTGCTTGGATCTCTT 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3348 GACTGCCATGAAAACTCCGAGAACTCCGAGAGACTGCGAGAGAGGCCCTTGCTTGGATCTCTTGAGCCTGTC 3407

Qy  2401 A 2401

Db  3408 A 3408
```

Search completed: April 4, 2004, 18:20:16
Job time : 656.552 secs

```
                    Copyright    GenCore version 5.1.6
                    Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     April 4, 2004, 16:58:13 ; Search time 116.44 Seconds
                                       (without alignments)
                                       11443.127 Million cell updates/sec

Title:           US-09-845-416-9_COPY_600_3000
Perfect score:   2401
Sequence:        1 acatgcattcaacatcgca.........tggatctcttgagcctgtca 2401

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        682709 seqs, 277745446 residues

Total number of hits satisfying chosen parameters:     1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued_Patents_NA:*
            1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1335.6	55.6	5952	4	US-09-687-875A-1	Sequence 1, Appli
2	1335.6	55.6	13977	4	US-09-488-970B-60	Sequence 60, Appl
3	1143.6	47.6	19307	3	US-08-836-022A-10	Sequence 10, Appl
4	1143.6	47.6	19307	3	US-09-427-048A-10	Sequence 10, Appl
5	487	20.3	6045	4	US-09-091-501B-7	Sequence 7, Appli
6	487	20.3	10320	4	US-09-091-501B-9	Sequence 9, Appli
7	401	16.7	3915	4	US-09-976-594-93	Sequence 93, Appl
8	79.4	3.3	200	4	US-09-091-501B-5	Sequence 5, Appli
9	78.6	3.3	200	4	US-09-091-501B-4	Sequence 4, Appli
10	78.6	3.3	200	4	US-09-091-501B-6	Sequence 6, Appli
11	76.6	3.2	7218	1	US-08-233-463-14	Sequence 14, Appl
12	63.6	2.6	238	4	US-09-687-875A-13	Sequence 13, Appl
13	46.2	1.9	505	4	US-09-621-976-15639	Sequence 15639, A
14	44.2	1.8	2574	4	US-09-666-313A-10	Sequence 10, Appl
15	44	1.8	1230025	4	US-09-198-452A-1	Sequence 1, Appli
16	43.4	1.8	1179	4	US-09-107-532A-1186	Sequence 1186, Ap
17	42.8	1.8	1690	4	US-09-620-312D-69	Sequence 69, Appl
18	42.8	1.8	7812	3	US-09-368-590-1	Sequence 1, Appli
19	40.4	1.7	832	4	US-09-621-976-2813	Sequence 2813, Ap
20	40.4	1.7	2223	1	US-08-257-073-4	Sequence 4, Appli
21	39.6	1.6	6960	2	US-08-841-349-3	Sequence 3, Appli
22	39.6	1.6	6960	4	US-09-431-184A-3	Sequence 3, Appli
23	39.6	1.6	8176	2	US-08-841-349-5	Sequence 5, Appli
24	39.6	1.6	8176	4	US-09-431-184A-5	Sequence 5, Appli
25	39.2	1.6	16995	4	US-09-961-527-82	Sequence 82, Appl
26	38.6	1.6	1751	4	US-09-620-312D-847	Sequence 847, App
27	38.6	1.6	1995	1	US-08-425-069-3	Sequence 3, Appli
28	38.6	1.6	1995	2	US-08-317-844B-3	Sequence 3, Appli
29	38.4	1.6	7672	4	US-09-220-132-24	Sequence 24, Appl
30	38.2	1.6	428	4	US-09-668-313A-3	Sequence 3, Appli
31	38.2	1.6	1131	6	5180810-3	Patent No. 5180810
32	38.2	1.6	1784	6	5180810-2	Patent No. 5180810
33	38.2	1.6	4439	4	US-09-668-313A-17	Sequence 17, Appl
34	38	1.6	1394	4	US-09-247-155-76	Sequence 76, Appl
35 c	36.6	1.5	1886	6	521083-1	Patent No. 5210183
36	36.6	1.5	1845	6	US-08-887-534A-22	Sequence 22, Appl
37	36.6	1.5	1845	4	US-09-527-431-22	Sequence 22, Appl
38	36.6	1.5	7075	4	US-08-956-171E-263	Sequence 263, App
39	36.2	1.5	289	3	US-09-007-005-17	Sequence 17, Appl
40 c	36.2	1.5	289	3	US-09-244-796-17	Sequence 17, Appl
41	36.2	1.5	713	4	US-09-091-501B-2	Sequence 2, Appli
42	36.2	1.5	2447	2	US-09-014-969-14	Sequence 14, Appl
43 c	36.2	1.5	11049	4	US-10-204-708-23	Sequence 23, Appl
44	36.2	1.5	168575	4	US-09-426-290-1	Sequence 1, Appli
45 c	36	1.5	724	4	US-08-956-171E-832	Sequence 832, App

ALIGNMENTS

```
RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc_feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match             55.6%;  Score 1335.6;  DB 4;  Length 5952;
Best Local Similarity   98.3%;  Pred. No. 0;
Matches 1350;  Conservative 0;  Mismatches 24;  Indels 0;  Gaps 0;

Qy       1028 TGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTACAAGACACTCATAGA 1087
              ||||| ||| || |||||||||||||||||| ||||||| |  |||||||||||||||
Db       2902 TGGAGAAGCATTCATAAAGGGTGAGTGAGCGAGAGCTGCTTGAAGAAACTCATAGA 2961

Qy       1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAGTTCTTGCCTGGCTTACAGAAGCTGAA 1147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2962 TTACTGCAACAGTTCCCCTGGACCTGGAAAGTTCTTGCCTGGCTTACAGAAGCTGAA 3021

Qy       1148 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 1207
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3022 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAG 3081

Qy       1208 GGAGTAAAGAGCTGATGAAACAATGGCAGAGACCTCCAAGGTGAAATTGAAGCTCACACA 1267
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3082 GGAGTAAAGAGCTGATGAAACAATGGCAGAGACCTCCAAGGTGAAATTGAAGCTCACACA 3141

Qy       1268 GATGTTTATCACAACCTGGATGAAAAATCCTGAGATCCCTGGAAGTTCC 1327
              ||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 2
US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
;   APPLICANT: Jones, Karen A.
;   APPLICANT: Volkmuth, Wayne
;   APPLICANT: Walker, Michael G.
;   TITLE OF INVENTION: BONE REMODELING GENES
;   FILE REFERENCE: PB-0014 US
;   CURRENT APPLICATION NUMBER: US/09/484,970B
;   CURRENT FILING DATE: 2000-01-18
;   NUMBER OF SEQ ID NOS: 172
;   SOFTWARE: PERL Program
; SEQ ID NO 60
;   LENGTH: 13977
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
;   NAME/KEY: unsure
;   LOCATION: 11721-11761, 12294, 13969
;   OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match   55.6%; Score 1335.6; DB 4; Length 13977;
Best Local Similarity   98.3%; Pred. No. 0;
Matches 1350; Conservative   0; Mismatches   24; Indels   0; Gaps   0;

Qy   1028 TGGGCAAACATCTGTAGATGGACAGAAGACCGTGGGTTCTTTTACAGACACTCATAGA 1087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8216 TGGAGAAGCATTCATAAAGGGTGAGTGAGGAGGCTGCTTTGGAGAAACTCATAGA 8275

Qy   1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAGTTCTGCCTGGCTTACAGAAGCTGAA 1147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8276 TTACTGCAACAGTTCCCCTGGACCTGGAAAGTTCTGCCTGGCTTACAGAAGCTGAA 8335

Qy   1148 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGAAAGGCTCCGTAAGGAAGACTCCAAG 1207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8336 ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAGACTCCAAG 8395

Qy   1208 GGAGTAAAAGACTGATGAAACAATGGCAGAAGACTCCAAGGTGAAAATTGAAGCTCACACA 1267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8396 GGAGTAAAAGACTGATGAAACAATGGCAAGAACCTCCAAGGTGAAAATTGAAGCTCACACA 8455

Qy   1268 GATGTTATCACAACCTGATGAAGACAGCCAAAAAATCCTGGAAGGTTCC 1327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8456 GATGTTTATCACAACCTGATGAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 8515

Qy   1328 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT 1387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8516 GATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT 8575

Qy   1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGGGT 1447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8576 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT 8635

Qy   1448 CTGCACCTTTCTCTCTGCAGGAACTTCTGGTGTGGTGGCTACAGCTGAAAGATGAATTAAGC 1507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8636 CTGCACCTTTCTCTCTGCAGGAACTTCTGGTGTGGTGGCTACAGCTGAAAGATGAATTAAGC 8695

Qy   1508 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGG 1567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8696 CGGCAGGCACCTATTGGAGGCGACTTTCAGCAGTTCAGAAGCAGAACGATGTACATAGG 8755

Qy   1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTCTTTGAGACTGTA 1627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8756 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTTCTTTGAGACTGTA 8815

Qy   1628 CGAATATTTCTGACAGAGAGCCCTTTGGAAGGACTAGAGAAAACTCTACCAGGACGCCAGA 1687

```
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE:  PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/08/836,022A
;    FILING DATE:
;    CLASSIFICATION:  435
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 08/331,381
;    FILING DATE:  28-OCT-1994
;  ATTORNEY/AGENT INFORMATION:
;    NAME:  Bak, Mary E.
;    REGISTRATION NUMBER:  31,215
;    REFERENCE/DOCKET NUMBER:  GNVPN.008PCT
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  215-540-9200
;    TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19307 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
US-08-836-022A-10

Query Match   47.6%;  Score 1143.6;  DB 3;  Length 19307;
Best Local Similarity   89.5%;  Pred. No. 0;
Matches 1230;  Conservative   0;  Mismatches 144;  Indels   0;  Gaps   0;

Qy  1028 TGGGCAAACATCTGTAGATGACGAGAGACCGCTGGGTTCTTTACAGACACTCATAGA 1087
         ||||| ||||||| ||| ||||| ||  ||||| |||||  ||||||||| |||||||
Db  6478 TGGGAAAACATTCATAAAGAGTAAGTGAGCAAGAGGCTGCTTGGAAGAAACTCATAGA 6419

Qy  1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGTGAA 1147
         ||||||||||||||| |||||||||||||| |||||| |||| |||||||||| |||
Db  6418 TTACTGCAGCAGTTCCCTGGACCTGGAGAAGTTTCTTTCCTGGATTACGGAGAGCAGAA 6359

Qy  1148 ACAACTGCCAATGTCCACAGATGCTACCCGTAAGGAAAAGGCTCCTAGAGACTCCAAG 1207
         ||||||||||||||||||||||||| ||| |||||||||||||| ||||||||||||
Db  6358 ACAACTGCCAATGTCCACAGAGCGCTTCCGTAAGGAAGAGCTCCTAGAGACTCCCAGG 6299

Qy  1208 GGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA 1267
         |||||||| ||||||||||||| |||||||||| ||||||||| |||||||||||||||
Db  6298 GGAGTCAGAGAGCTGATGAAGACCAAGATCCTCCAAGGAGAAATTGAAACTCACACA 6239

Qy  1268 GATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGAAGGTTCC 1327
         |||| |||||||| |||||| |||||| ||||||||||||| |||||||||||||||||
Db  6238 GATATCTATCACAATCTTGATGAAAAATCGCCAAAAAATCCTGAGATCCCTGAAGGTTCG 6179

Qy  1328 GATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAGTGAACTT 1387
         ||||||||||||||||| ||||||||||||||| ||||||| ||||||||||||||||
Db  6178 GATGAAGCACCCCTGTTACAAAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAACTT 6119

Qy  1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT 1447
         |||||||||||||||||| ||||||||||||||||||| |||||||||||||||||||||
Db  6118 CAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAAGTTCTGACCAGTGGAAGCGT 6059

Qy  1448 CTGACCTTTCTTCTGCAGGAACTTCTGTGTGCTACAGCTGAAAGATGATGAATTAAGC 1507
         ||| |||||||||||||||||||||| |||| ||||||||||||||||||||||||||
Db  6058 TTGCATCTTTCTCTTCAGGAACTTCTGTTTGGCTACACGCTGAAAGATGATGAACTGAGC 5999

Qy  1508 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAGGCAGAGGCAGATGTACATAGG 1567
         |||||||||||| ||||||||||||| |||| |||||||||||||||||| ||||||||||
Db  5998 CGTCAGGCACCCATCGGTGGTGATTTCCCAGCAGTTCAGAGGCAGAGGCAGATGATATACATAGG 5939

Qy  1568 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCACATGAGTACTCCTTGAGACTGTA 1627
         ||||||||||||||||||||||| |||||| |||||||||||| ||||||| ||||||||||
Db  5938 GCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTCTGAGACTGTG 5879

Qy  1628 CGAATATTTCTGACAGAGCCAGCAGCCTTTGGAAGAACTAGAGAAACTCTACCAGGAGCCCAGA 1687
         || ||||||||||||||| |||||||||||||||||||| || || ||||||||||||||||||
Db  5878 AGAATATTTCTGACAGAGCCAGCAGCCTTTGGAAGAACTAGAGAAACTCTACCAGGAGCCCAGA 5819
```

```
Db  8816 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         CGAATATTTCTGACAGAGCCAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGA 8875
Qy  1688 GAGGTGCCTCCTGAGGAGGAGCCCAGAATGTCACTCGGCTTCACTCGGCTTCACGAAAGCAGGCTGAG 1747
Db  8876 GAGCTGCCTCCTGAGGAGGAGCCCAGAATGTCACTCGGCTTCACTCGGCTTCACGAAAGCAGGCTGAG 8935
Qy  1748 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTTCGCTGACTTCGCTGACTTGGCAGAGAAAATA 1807
Db  8936 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTTCGCTGACTTCGCTGACTTGGCAGAGAAAATA 8995
Qy  1808 GATGAGACCCCTGAAAGACTTCAAGAACTTCAAGAGGGCCACGATGAGCTGGACTTCAAG 1867
Db  8996 GATGAGACCCCTGAAAGACTTCAAGAACTTCAAGAGGCCACGATGAGCTGGACTTCAAG 9055
Qy  1868 CTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGATCCTCCTCATTGAC 1927
Db  9056 CTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGCGATCTCCTCATTGAC 9115
Qy  1928 TCTCTCCAAGATCACCTCGAGAAGTCACTTCCAGGCGACACTTCCAGGACATTCGGCCTCTGAAA 1987
Db  9116 TCTCTCCAAGATCACCTCGAGAAGTCAAGGCACTTCGAGGACAATTGCGCCCTCTGAAA 9175
Qy  1988 GAGAACGTGAGCCACGTCAATGCACTCTCGCCGCAGTTACCACTTTGGGCATTCAGCTC 2047
Db  9176 GAGAACGTGAGCCACGTCAATGCACTCTCGCCGCAGCTTACCACTTTGGGCATTCAGCTC 9235
Qy  2048 TCACCGTATAAACTTCAGCACTCTGGAAGACCTGAACACCAGATGGAGGAGCTTCTGCAGGTG 2107
Db  9236 TCACCGTATAAACTTCAGCACTCTGGAAGACCTGAACACCAGATGGAGGAGCTTCTGCAGGTG 9295
Qy  2108 GCCGTCGAGGACCGAGTCAGGCAGCTGAGCCACGGGACTTTGGTCCAGCATCT 2167
Db  9296 GCCGTCGAGGACCGAGTCAGGCAGCTGAGCCACGGGACTTTGGTCCAGCATCT 9355
Qy  2168 CAGCACTTTCTTCCACGTCGTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGGCCAAACAAA 2227
Db  9356 CAGCACTTTCTTCCACGTCGTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGGCCAAACAAA 9415
Qy  2228 GTGCCCTACTATATCAACCACGAGACTTGCTGGGACCATCCCAAAATGACA 2287
Db  9416 GTGCCCTACTATATCAACCACGAGACTTGCTGGGACCATCCCAAAATGACA 9475
Qy  2288 GAGCTTCACCAGTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTCAGCTTATAGGACTGCC 2347
Db  9476 GAGCTTCACCAGTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTCAGCTTATAGGACTGCC 9535
Qy  2348 ATGAAACTCCGAAGACTCGCGAAGAGGCCCTTGCTTGCTTCTTCTTGAGATCTCTTGTCA 2401
Db  9536 ATGAAACTCCGAAGACTCGCGAAGAGGCCCTTGCTTGCTTCTTCTTGAGATCTCTTGTCA 9589
```

```
RESULT 3
US-08-836-022A-10//C
; Sequence 10, Application US/0836022A
; Patent No. 6001557
; GENERAL INFORMATION:
;    APPLICANT:  Trustees of the University of Pennsylvania
;    APPLICANT:  Wilson, James M.
;    APPLICANT:  Fisher, Krishna J.
;    APPLICANT:  Chen, Shu-Jen
;    APPLICANT:  Weitzman, Matthew
;    TITLE OF INVENTION:  Improved Adenovirus Virus and
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, P O Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

```
Query Match      47.6%;   Score 1143.6;   DB 3;   Length 19307;
Best Local Similarity  89.5%;   Pred. No. 0;
Matches 1230;  Conservative  0;  Mismatches  144;  Indels  0;  Gaps  0;

Qy   1028  TGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAGACACTCATAGA  1087
                 |||| ||||||||||| || |||||||||||||||  ||| ||| || ||||| || |||
Db   6478  TGGGGAAACATTCATAAAAGAGTAAGTGAGCAAGAGGTGCTTGGAGAAAACTCATAGA  6419

Qy   1088  TTACTCGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGAAGCTGAA  1147
                 |||||| | |||||||||| |||||||||||||| |||| |||||  |||||||||||||
Db   6418  TTACTCGCAGCAGTTCCCCTGGACCTGGAAAAGTTCTTTCCTGGATTACGGGAAGCAGAA  6359

Qy   1148  ACAACTGCCAATGTCCTACAGATGCTCACCCGTAAGGAAGGCTCCGTAAGGAAGACTCCAAG  1207
                 |||||||||||||||| ||||| |||||||| |||  ||| |||||||||| |||||||||
Db   6358  ACAACTGCCAATGTCCTACAGACGCTTCCCGTAAGGAAGGCTTCCGTAAGGAAGACTCCAGG  6299

Qy   1208  GGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGCTCACACA  1267
                 |||||||||||||||||||||||| |||||||||| ||||||||||| |||||||||||
Db   6298  GGAGTCAGAGAGCTGATGAGAGCTGATGAAACAATGGCAAGATCCTCCAAGGAGAAATTGAAAGCTCACACA  6239

Qy   1268  GATGTTTATCACAACTGGATGAAACACGCCAAAAAATCCTGAAGATCCTGGAAGGTTCC  1327
                 |||| |||||||| ||| ||| ||||  |||||||||||||| ||||| ||| ||||||||
Db   6238  GATATCTATCACAATCTTGATGAAAATGGCAAAAAATCCTGAAGATCCTGGAAGGTTCG  6179

Qy   1328  GATGATGCAGTCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTT  1387
                 |||||||||||||||||||||||||||||||||| ||||| |||||||||||||||||||| ||||
Db   6178  GATGAAGCAGCACCCGTTACAAAAGACGTTTGGATAACATGAATTTCAAGTGGAGTGAACTT  6119

Qy   1388  CGGAAAAAGTCTCTCAACATTAGGTCCATTTGGAGCCAGTTCTGACCAGTGGAAGCGT  1447
                 |||||||||||||||||| ||||| | |||||||||||||||||||| ||| ||||||||||
Db   6118  CAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGCCAGTTCTGACCAGTGGAAGCGT  6059

Qy   1448  CTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAGAAGAATGAATTAGC  1507
                 ||  ||||||||||||||||||||||||||||||||| ||||||||||| |||||||||||
Db   6058  TTGCATCTTTCTCTTCTTCAGGAACTTCTTGTTGTTGGCTACAGCTGAGAAGAATGAACTGAGC  5999

Qy   1508  CGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAGGAGCAGAAGCAGCATGTACATAGG  1567
                 ||| |||||||| ||||||  ||||||||||||||||||||| ||||||||||||||||||||||
Db   5998  CGGCAGGCACCCATCGGTGGTGATTTCCCAGCAGTTCAGAAGAAGCAGCAGAATGATATACATAGG  5939

Qy   1568  GCCTTCAAGGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA  1627
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5938  GCCTTCAAGGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTG  5879

Qy   1628  CGAATATTTCGACAGAGCAGCCTTTGGAGGAGGACTAGAAAACTCTACCAGGAGCCCAGA  1687
                 |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5878  AGAATATTTCGACAGAGCAGCCTTTGGAGGAGGACTAGAAAACTCTACCAGGAGCCCAGA  5819
```

```
Qy   1688  GAGCTGCCTCCTGAGGAGAGAGAGCCCAGAATGTCACTCGGCTTCACGAAAGACAGGCTGAG  1747
                 |||||| ||||||||||||||||||||||||||||||||||  |||| | ||||||||||||
Db   5818  GAACTGCCTCCTGAGGAAAGAGAGCTCACTCGGCTCCTACGAAAGACAGGCTGAA  5759

Qy   1748  GAGGTCAATACTGAGTGGGAAAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAATA  1807
                 ||||| ||||||||||| |||||||||||||||||||| ||||||||||||||||||||||
Db   5758  GAGGTCAACGCTGAATGGGACAAATTGAACCTGATTGGCAGACTGATTGGGCAGAGAAAATA  5699

Qy   1808  GATGAGACCCTGAAAGACTCCAGGAACTTCAAGACTGAGCTGACCCTCAAG  1867
                 ||||||||| |||||||||||||||||||||||||||||| || |||||||||
Db   5698  GATGAAGCTCTTGAAAGACTCCAGGAACTTCAGGAAGTCGCCGATGAACTGACCCTCAAG  5639

Qy   1868  CTGCCCAAGTCAGTGAGTGATCAAGGATCCTGGCAGCCGGTGCCGCAGCTCCTCATTGAC  1927
                 |||||| |||||||||| ||||||||||||||| ||||||||||| |||| ||||||||||
Db   5638  TTGGGCCAAGTCAGCTTGAGTGATCAAGGGATCTGGGCAGCCAGTGGGGGATCTCCTCATTGAC  5579

Qy   1928  TCTCCCAAGATCACCTCGAGAAAGCACTTCGAGGAGAAATTGCGCCTCTGAAA  1987
                 |||||| ||||||||||| ||||||||||||||| |||||||||||||||||||
Db   5578  TCTCTGCAAGATCACCTTGAAAAAGTCAGGCACTTCGGGGGAGAAATTGCACCTCCTTAAA  5519

Qy   1988  GAGAAGTGAGCCACGTCAATGACCTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTC  2047
                 |||||||||||||||  |||| |||||||| ||| ||||||||||||||||||||||||||
Db   5518  GAGAATGTCAATCGTGTCAATGACCTTCGCCAGCATCAGCTTGACACACTGGGCATTCAGCTC  5459

Qy   2048  TCACCGTATAACCTCTGCGAAGACCTGAACACCAGAATGGAAGTTCTGCAGGTG  2107
                 |||||| ||||| |||||| ||||||||||| ||||| ||||| |||||||||||||
Db   5458  TCACCTTATAACCTCAGCACTTGGAAGATCTGAATACCAGAATGGAAGTTCTCTACCAGGTG  5399

Qy   2108  GCCGTCGGAGGACCGGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCT  2167
                 |||| |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||
Db   5398  GCTGTGGAGGACCGTGTCAGACAGCTGCATGAAGCCCACAGGGACTTTGGTCCTGCATCC  5339

Qy   2168  CAGCACTTCTTCCCACGTCTGTCCAGGGGTCCCTGGGAGAGGAGCCATCCGCCAAACAAA  2227
                 |||||||||||| |||||| ||||| |||||||||||||||||||||||| || ||||||||
Db   5338  CAGCACTTCTTCCCACTTCAGTTCCAGGGTTCCCTGGGAGAGGAGCCATCACCAAAACAAA  5279

Qy   2228  GTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACA  2287
                 ||||||||||||||||||||||||||||| ||||| ||||||||||||||||||| |||||
Db   5278  GTGCCCTACTATATCAACCACGAGACCCAAAACCACTTGTGGGACCACCCAAAATGACA  5219

Qy   2288  GAGCTCTACCAGTCTTTAGTGACCTGAATAAATGTCAGATTCTCAGCTTATAGGACTGCC  2347
                 ||||| ||||||||||||||||||||||| |||||||||||||| | |||| ||||||||||
Db   5218  GAGCTCTACCAGTCTTTAGTGACCTGAATAAATGTCAGGTTCTCCGCGTATAGGACTGCC  5159

Qy   2348  ATGAAACTCCGAAGACTGCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCA  2401
                 |||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   5158  ATGAAGCTCAGAAGGCTCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCA  5105
```

RESULT 4
US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
Patent No. 6203975
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
Wilson, James M.
Fisher, Krishna J.
Chen, Shu-Jen
Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
Methods of Use Thereof
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match     20.3%;   Score 487; DB 4; Length 6045;
Best Local Similarity   58.3%;   Pred. No. 1.2e-139;
Matches  878;   Conservative   0;   Mismatches  620;   Indels   9;   Gaps   1;

Qy    894 TGCTTCAAGAAGATCTAGAAGACAAGAACAAGTCAGGTCAATTCTCACTCACATGGTGG  953
          ||||         ||| |                  ||       |||   |
Db   2871 TACTTGCTTCAAGAACTGGGTCCTGGAGATGGTATCGGTCGGTTCGATAACGTCCTGC 2930

Qy    954 TGGTAGTTGATGAAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGG 1013
          |                                  ||||||
Db   2931 AGAAACTCCTGGAGGAATATGGGAGTGATGACAAGGAATGTGAAAGAAAACCACAGAGT 2990

Qy   1014 TATTGGGAGATCGATGGGCAAACATCTGTAGATGACAGAGAGACCCGTGGGTTCTTTTAC 1073
          ||||        |    |           ||||||
Db   2991 ACTTAAAAACATCATGATCCAAACATCAAAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGC 3050

Qy   1074 AAGACACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGC 1133
          ||||         ||||
Db   3051 AGGCTGAGTGGAGGACCGGTGCGAGCCTGCTTCAGAGGATCTGGAAAACTTCCTGAAGTGGA 3110

Qy   1134 TTACAGAAGCTGAAACAACTGCCAATGTCCTACAGATGTCCTAAGGAAAGGCTCC 1193
          ||            |||||
Db   3111 TCCAAGAAGCAGAGACCACCAGTGAATGTGCTTGTGGATGCCTCTCATCGGGAGAAATGCTC 3170

Qy   1194 TAGAAGACTCCAAGGAGTAAAAGAGCTGATGAAACAATGCGATGAACAAGGTGAAA 1253
          ||            ||
Db   3171 TTCAGGATAGTATCTTCGGCCAGGGAACTCAAACAGCCAGATGCGGAGCATCCAGGCAGAAA 3230

Qy   1254 TTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGAT 1313
          ||                |    ||
Db   3231 TTGATGCCCACAATGACATATTTAAAAGCATTGACGGAAACAGGCAGAGATGGTAAAAG 3290

Qy   1314 CCCTGGAAGGTTCCGATGATGGCAGCAGTCGTGTTACAAAGACGTTTGGATAACATGACTTCA 1373
          ||||         ||
Db   3291 CTTTGGGAAATTCTGAAGAGTGCTACTATGCTTCAACATCGACTGACTGATGATGAACCAAA 3350

Qy   1374 AGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGTCTCAACATTGGGAGCCAGTTCTG 1433
          ||||
Db   3351 GATGGAATGACTTAAAAGCAAAATCTGCTAGCATCAGCGGGGCCATTGGGAGGGCCAGCGGCTG 3410

Qy   1434 ACCAGTGGAAGCGTCTGCACCTTTCTGCGAGGAGGACTTCGGTGTGGCTACAGCTGAAAG 1493
          ||||
Db   3411 AGAAGTGGAACAGGTTCGCTGATGTCCTTAGAAGAACTGATCAAATGGCTCAAATATGAAAG 3470

Qy   1494 ATGATGATTAAGCCGGCAGGCAGCCGCACTTTCCAGTGTCAGCAGCAGA 1553
          ||||
Db   3471 ATGAAGAGCTTAAGAACAAATGCCTATTGGAGAGATGTTCCAGCCTTACAGCTCCAGT 3530

Qy   1554 ACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTA 1613
          ||||
Db   3531 ATGACCATTGTAAGGCCCTGACAGCGGGAGTTAAAGGAGAAGAATATTCTGTCCTGAATG 3590

Qy   1614 CTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAAC--- 1670
          ||||
Db   3591 CTGTCGACCAGGCCCGAGTTTTCTTGGCTGATCAGCCAATTGAGGGCGCCCTGAAGAGCCAA 3650

Qy   1671 -----TCTACCAGGAGCCAGAGAGCTGCCTCCTGAGGAGGAGCCCAGAAGTGTCACTC 1724
          ||       |
Db   3651 GAAGAAACCTACAGAATCAAATCAAAAACAGAATTAACCTCCTGAGGAGGAGCCCAAAAAGATTGCCA 3710
```

```
Qy  1688 GAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1747
         ||         ||||||||
Db  5818 GAACTGCCTCCTGAAGAGAAGAGAGCTCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAA 5759

Qy  1748 GAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACCTCCGCGACTGGCAGGAAAAATA 1807
         |||||||||                 ||||||
Db  5758 GAGGTCAACGCTGAATGGGACACAAATTGAACCTTGCGCTCAGCTGATTGGCAGGAGAAAATA 5699

Qy  1808 GATGAGACCCTTGAAAGACTCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAG 1867
         ||||    ||| | ||||||
Db  5698 GATGAAGCTTCCAGGAACTTCAAGGAAGCTCAGGAGTCTGCCGATGAACTGGACCTCAAG 5639

Qy  1868 CTGCGCCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGTGGGGCCGATCTCCTCATTGAC 1927
         ||         |||||||
Db  5638 TTGCGCCAAGCTGAGGTGATCAAGGATCCTGCAGCCAGTGGGGGATCTCCTCATTGAC 5579

Qy  1928 TCTCTCCAAGATCACCTCGAGAGAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAA 1987
         ||||         ||||||
Db  5578 TCTCTGCAAGATCACCTTGAAAAAGTCAAGGCACTTCGGGGAGAAATTGCACCCTCTTAAA 5519

Qy  1988 GAGAACGTGAGCCACGTCAATCACCTTGCTCGCCAGCTTACCACTTGGGCCATTCAGCTC 2047
         ||||
Db  5518 GAGAATGTCAATCGTCAATGACCTTGCACATCAGCTGACCACACTGGGCCATTCAGCTC 5459

Qy  2048 TCACCGTATAACCTCAGGCACTCTGGAGACCACCAGATGGAAGCTTCTCCAGGTG 2107
         ||||
Db  5458 TCACCTTATAACCTCAGCACTTTGGAAGATCTGAATACCAGGAGGCTTCTACAGGGTG 5399

Qy  2108 GCCGTCGAGGACCGAGTCAGGCAGCTGCAGGGTCCTGGGGAGGAGACGACTTTGGTCCAGCATCT 2167
         ||||
Db  5398 GCTGTGGGAGACCGTGTCAGACAGCTGCATGAAGTGCATGGGGACTTGGGTCCTTGCCATCC 5339

Qy  2168 CAGCACTTTCTTTCCACGTCTGTCCAGGGTCCTGGGAGGAGAGCCATCTCGCCAAACAA 2227
         ||||
Db  5338 CAGCACTTTCCCTTTCCACTTCAGTTGACGACTCAGGGTCCTGGGAGGAGAGCCATCTCCACCAAACAA 5279

Qy  2228 GTGCCCTACTATATCAACCACGACTCAAACAACTGCTGGGACCATCCCAAAATGACA 2287
         ||||
Db  5278 GTGCCCTACTATATCAACCACGACCACGAGACCCACACCTTGTTGTGGGACCATCCCAAAATGACA 5219

Qy  2288 GAGCTCTACCAGTCTTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 2347
         ||||
Db  5218 GAGCTCTACCAGTCTTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCC 5159

Qy  2348 ATGAAACTCCGAAGACTGCAGAAGGCTCAGAAGGCCTTGCTTGGATCTCTTGAGCCTGTCA 2401
         ||         ||||||||
Db  5158 ATGAAGCTCAGAAGACTCAGAAGGGCTCCAGAAGGCCCCTTTGCTTGGATCTCTTGAGCCTGTCA 5105
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: (11)..(10312)
;   FEATURE:
;     OTHER INFORMATION: Description of Artificial Sequence: Full length
;     OTHER INFORMATION: utrophin construct
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (724)..(758)
;     OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match          20.3%;  Score 487;  DB 4;  Length 10320;
Best Local Similarity 58.3%;  Pred. No. 1.8e-139;
Matches  878; Conservative   0; Mismatches 620; Indels   9; Gaps   1;

Qy    894 TGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACATGGTGG    953
          ||| |||||| |  | |||  |  ||| ||||||  ||   |||||   ||   |   |
Db   7146 TACTGCTTCAAGAACTGGGTCCTGGAGATGGTATCGTCATGGCGTTCGATAACGTCTCTGC 7205

Qy    954 TGGTAGTTGATGAATCTAGTGGAGATCACGCAACTCGTGCTTGCTTGGAAGAACAACTTAAGG 1013
          |  ||  | | | ||  | ||| |  |||  |  |||| |  | | |||||| |  ||  |
Db   7206 AGAAACTCCTGGAGGAAATGGGAGTGATGACAACAAGGAATGTGAAAGAAACCACAGAGT   7265

Qy   1014 TATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGGCTGGGTTCTTTTAC   1073
          | ||  | ||| ||    | |||||  | | |||||||||| || || || |||| | ||
Db   7266 ACTTAAAAACATCATGGATCAATCTCAAACAAAGTATTGCTGACAGACGACGCCTTGG     7325

Qy   1074 AAGACACTCATAGATTACTGCAACAGTCCCCTGGACCTGGAAAAGTTTTCTTGCTGGC     1133
          | | |  | | ||||| | |  ||||| || ||||| || |||| |||| ||| |||
Db   7326 AGGCTGAGTGGAGGACGGTGCAGCCTCTGCAGAGATCTGGAAAAACTTCCTGAAGTGGA    7385

Qy   1134 TTACAGAAGCTGAAACAACTGCCAATGTCCATCCTACAGGATGCCTGAAGGAAAAGGCTCC  1193
          | ||||| |||||||| |  |||||||  |||||  |||||||||   | |||||   |||
Db   7386 TCCAAGAGCAGAGACCCAGTGAATGTGCTTGTGGATGCCTTCTCATCGGGAGAAATGCTC   7445

Qy   1194 TAGAGAACTTCAAGGAGTAAAAGAAGAGCTGATGAAACAATGGCAGACCTTCAAGGTGAAA  1253
          | ||| || ||||||||  |||| ||||||||||| | |  ||| |  |||||   |||||
Db   7446 TTCAGGATAGTATCTTGGCCAGGAAGACAAATCAGCAGGACATCCAGGCAGAAA        7505

Qy   1254 TTGAAGCTCACACAGATGTTATCACAACCTGATGAAAACAGCCAAAAATCCTGAGAT      1313
          | |||| | || ||| |||| |  | |  |||||||||||||| |||| | ||||
Db   7506 ATGAAGAGCTTAAGAACAAAATGCCTATTGGAGGAGGATGTTCCAGCCTTACAGCTCCAGT  7565

Qy   1314 CCCTGGAAGGTTCCGATGATGCAGTCGTCCTGTTACAAGACGTTTGGGATAACATGAACTTCA 1373
          |  ||||| | |||||||||  |||| || |  |||| | |||  |||||||||| |||||
Db   7566 CTTTGGAAATTCTGAAGAGTCTGAAGAAATCTGAAGCATCGAACATCGAACGACTGGATGATGAACCAAA 7625

Qy   1374 AGTGGAGTGAACTTCGGAAAAAGTCTCTCCAACATTAGTCCCATTGGAAGCCAGTTCTG    1433
          | |||| |||| || ||||||||| | |  ||||| ||  |||||||| ||||| |||
Db   7626 GATGGAAATGACTTAAAAGCAAAATCTGCTAGCATCAGGGCCCCATTTGGAGCCAGCCGCCTG 7685

Qy   1434 ACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGTGGTGTGGCTACAGCTGAAAAG 1493
          | |||||||| ||| | |||   | ||  ||||| | ||||  |  || | ||||  |||||
Db   7686 AGAAGTGGAACAGGTTGCTGATGTCCTTAGAGAGAACTGATCAAATGCTGAATATGAAAG     7745

Qy   1494 ATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAGCAGA    1553
          ||||| ||||||||| || || |  |||||||||  | ||||||| || |||||||
Db   7746 ATGAAGAGCTTAAGAACAAAATGCCTATTGGAGGAGGATGTTCCAGCCTTACAGCTCCAGT  7805

Qy   1554 ACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTA    1613
          | ||| | || || ||| | || | |||||||||| |||| ||| |  ||| |||| |
Db   7806 ATGACCATTGTAAGGCCCTGAGACGCGGAGTTAAAGAGAAGAATATTCTGTCCTGAATG    7865

Qy   1614 CTCTTGACTGTACGAAATATTCTGAAGGACTAGAGAAAC---                    1670
          ||||||| || |||| ||||| |  |||  || ||||||
Db   7866 CTGTGACCAGGGCCCGAGTTTCTTGGCTGATCAGCCAGTTCAGCAGCTCTGAAGAGCCAA  7925

Qy   1671 ------TCTACCAGGAGCCACAATCAAAAACAGAATTAACTCCTGAGGAGAGCCAGAGCCCAGAAGTGTCACTC 1724
               |||| |||||||| | |||||||||| ||||||||||||||||||| |||  ||||| | | |  |
Db   7926 GAAGAAACCTACAATCAATCAAAAAACAGAATTAACTCCTGAGGAGAGAGAGCCCAAAAGATTGCCA      7985
```

```
Qy   1725 GGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCACT  1784
          |||  |  | |  |  |   |||| ||    ||  |  |  | ||||  |||||| |  |
Db   3711 AAGCCATGGCCGCAAACAGTCTTCTGAAGTCAAAGAAAAATTGGGAAGTCTAAATGCTGAA  3770

Qy   1785 CCGCTGACTGGCAGCAGAAAAATAGATGAGACCCTTGAAAGACTTCAGAGACTTCAGAGG   1844
          |  |||| |  |   | |||| ||| | ||||| | ||||| ||||||||| |||| |||
Db   3771 CTAGCAATTGCAAAGCAAGCAGCATTGGAGAAACTCAGAGACCTCCAGGGAG          3830

Qy   1845 CCACGGATGAGCTGGACCTCAAGCTGCGGCCAAGCTGAGGTGATCAAGGATCCTGGCAGC   1904
          || | ||||| |||||| | ||| || |    |||||| |   |||||| | | || |||
Db   3831 CTATGGATGACCTGGACGCTGGACATGAAGGAGGCAGAGTCCGTGCGGAATGGCTGGAAGC  3890

Qy   1905 CCGTGGGGCGATCTCCTCCATTGACTCTCCAAGATCACCTCGAGAAGTCAAGGCACTTC   1964
          ||||||| | | |  || |||||| | | |  ||| | ||  |||||  | ||||| |
Db   3891 CCGTGGGAGACTTACCATTGACTGACTGCGTCGCAGGATCACATTGAAAAAATCATGGCATTA 3950

Qy   1965 GAGGAGAAAATTGCGCTTCGTGAAAGAAGAACGTGAGCCACGTCAATGACCTGCTCGCCAGC  2024
          || || ||||| || | |  |  | ||  ||| |||  |  | ||||||| |  || ||||
Db   3951 GAGAGAAATTGCACCAATCAACTTAAAGTTAAAACGGTGAATGATTTATCCAGTCAGC     4010

Qy   2025 TTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAACA   2084
          |  |||| | | ||| ||| ||| |||| |||| |||||| ||| | |  |||| |||  |
Db   4011 TGTCTCCACTTGACCTGCCATCCCTCTCTAAAGATGTCTCGCCAGCTAGATGACCTTAATA   4070

Qy   2085 CCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCCGAGTCAGCCAGCTGCATGAAGCCC   2144
          |  |||||||  |||| |||| | | || || |||  |||||||||| || |  | |||||
Db   4071 TGCGATGGAAACTTTTTACAGGTTTCTGTGGATGATCGCCTTAAACAGCTTCAGGAAGCCC   4130

Qy   2145 ACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCTTCCACGTCGTGTCCAGGGTCCCTGGG  2204
          ||||| ||||| |||||| |||||||| |||||  | | ||| | ||  | ||  |  |||
Db   4131 ACAGAGATTTTTGGACCATCCTCTCAGCATTTCTCTCTACGTCCAGTCGCCGTGGC       4190

Qy   2205 AGAGGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACCAGGACTCAAACAACTT   2264
          | || ||||| | ||||| || ||||||||| |||||||||| ||    ||||||||
Db   4191 AAAGATCCATTTCACATAAACTAAAGTGCCCTATTACATCATCCACCATCAAACAACAGACCACCT 4250

Qy   2265 GCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTAGCTGACCTGAATAATGTCA    2324
          | ||||||||||| ||||||| || ||||| | |||| | ||||||||||||||| | |
Db   4251 GTTGGGACCATCCAAAATGACCGAACTCTTTCAATCCCTTGCTGACCTGAATAATGTAC    4310

Qy   2325 GATTCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAGGCCCTTTGCTTGG     2384
          | |||||| |   | || ||  || | ||  ||||||||| ||||  | |   | ||||
Db   4311 GTTTTTCTGCGCACCGTACGACGCAATCAATCGGAAGAACTCCGAAGACTACAAAAGCACTATGTTTGG 4370

Qy   2385 ATCTCTT  2391
          |||||||
Db   4371 ATCTCTT  4377
```

Query Match   16.7%;   Score 401; DB 4; Length 3915;
Best Local Similarity 62.5%; Pred. No. 3.8e-113;
Matches 644; Conservative   0; Mismatches 385; Indels   2; Gaps   1;

```
Qy  1725 GGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAACCTGCACT 1784
         ||||| |||| ||  |||||
Db  7986 AAGCCATGGCGCAAACAGCTCTTCTGAAGTCAAAGAAAAATGGGAAAGTCTAAATGCTGTAA 8045

Qy  1363 CATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGA 1422
         ||||||||||||||||| ||||||| ||||||||||||| ||||||  |||| ||||||
Db  544  CATGAATCTGTTGTTGAATGAATAAAAAAAAAAAGTCTCAAAACCTCCGGCCGCTCGCCTAGA 603

Qy  1785 CCGCTGACTGGCAGCGAGAAAAATAGATGAGACCCTGAAAGACTTCAGGAACTTCAAGAGG 1844
         ||||  |||||| |||||||| ||||||||
Db  8046 CTAGCAATTGGCAAAAGCAAGTGGACAAGGCATTGGACAAACTTCAGAGAAACTCAGCAGGGAG 8105

Qy  1423 AGCCAGTTCTGACCAGTGGAACGGTCTGGAACGGTCTGATCAAGGGGATCCTGCGAGC 1482
         ||||  ||||||||| ||||
Db  604  GGCTTCTCAGACCACCACAGTGGAAAGCTTCAGCTCCCTTCAAGAGATTATTGACTGGCT 663

Qy  1845 CCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGTGGATCAAGGGATCCTGCGCAGC 1904
         ||||||| |||||||| |||  |||||||||| |||||||||||
Db  8106 CTATGGATGACCTGGACGCTGACATGAAGCGAGGAGGCAGAGTCCGTGCCGAATGGCTGAAAGC 8165

Qy  1483 ACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGCGGACTTTCCAGCAGT 1542
         |||||||||||||||||  ||||||
Db  664  CAGCCAAAAGGATGAGGAGTTGTCAGCTCAGCTGGCCCCTACAGGGCGGATGGGCCCTGGT 723

Qy  1905 CCGTGGGGCGGATCTCCTCATTGACTCCCAAGATCACCTCTGAGAAAGTCAAGGCACTTC 1964
         |||||||| |||||||||||
Db  8166 CCGTGGGAGCTCACTGGACTTCAGTTCGCTGCAGATATCCGCTGCGAGATCACCATGGGCATTTA 8225

Qy  1543 TCAGAAGCAGAACGATGTACATAGGGGCTTCAGAGAGGGAATTGAAAACTAAAGAACCTGT 1602
         |||||||
Db  724  GCAACAGGAGAAGGAGGAACACCATGCGCGGCCTTTATGGAAGAAGTCAAGTCTCCGGGGCCCCTA 783

Qy  1965 GAGGAGAAATTGCGCCTCTGAAAGAAGACGTGCAGCCACGTCAATGACCTTGCTCCGCCAGC 2024
         ||||||||| ||||||||||| ||| ||
Db  8226 GAGAAGAAATTGACCAATCAACTTTAAAGTTAAACGGTAAAGTAATCCAGTCAGC 8285

Qy  1603 AATCATGAGTACTCTTGACAGTGTAGAAATATTTCTGACAGAGCAGCCCTTTGGAGGGACT 1662
         ||||||||||| ||||||| ||||
Db  784  CATCTATTCGTGTGGAGTCAGCTCAGGCCCTTCCTGTCCCAGCACCCATTTGAGGAGTTT 843

Qy  2025 TTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCTGGAAGACCTGAACA 2084
         ||||||||  |||
Db  8286 TGTCTCCACTTGACCTGCCATCCCCTCTCTCTAAAGATGTCTCGCCCAGCTGAGGAGTGGATGAC 8345

Qy  1663 AGAGAAACTTCTACCAGAGCCCAGAGCCGCTGCCTCCTGAGGAGAGAGCCAGAATGTCAC 1722
         |||||||| |||||||||||||
Db  844  AGAGGAGCCTCATTCTGAGAGCAAAAGATACCTCCCGAAACAGCCGGATCAGAATCTCAG 903

Qy  2085 CCAGATGGAAGCTTCTGCAGGTGGCCGTGCAGGACCAGCAGTCAGGCCCTCATGTCAGGAGCCC 2144
         || ||||
Db  8346 TGCCGATGGAAACTTTTACAGGTTTCTGTGGATGCATTTTCTGTTCAGGTCTCAGCTTCAGGAAGCCC 8405

Qy  1723 TCGGCTTCTACGAAAGCAGCCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAACCTGCA 1782
         || |
Db  904  CCGCTTTGTATGGAAGCAGGCGACGGTGGCCAGTGAACTGTTGGGAGAAGTTGACAGCCCG 963

Qy  2145 ACAGGGGACTTGGTCCAGCATCTCAGCACTTCTTCTTCCACGTCTGTCCAGGTCCCTGGG 2204
         |||||||| || |||||
Db  8406 ACAGAGATTTTGGACCCATCCTCTCTCAGCATTTCTTCTCTCTCTACGTCAGTCAGTCCAGGGGT 8465

Qy  1783 CTCCGCTGACTGGCAGCAGAGAAAAATAGATGAGACCCTGAAAGACTTCAGAGAACTTCAAGA 1842
         ||| |||
Db  964  CTGTGTGGACCAGCACCGTCACATTGAGCGGACTCGGTGAGCAGCTCTTGGAGATTCA--G 1021

Qy  2205 AGAGAGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCACCAGACTCAAACAACTT 2264
         ||||||||||||||| |||||||
Db  8466 AAAGATCCATTTCACATATAAAGTGCCCTATTACCATTTCAACCATCAACACACAGACCACCT 8525

Qy  1843 GGCCACGGATGAGCTGGACCTCAAGCTGCGCGCAAGCTGGAGGTGATCAAGGGATCCTGGCA 1902
         |||| |||||||||||
Db  1022 CGTCCGATGGAGGAACTAAGCACTACTCTCAGCCAAGCTGAGGAGTGAGGGAGCACTTGGGA 1081

Qy  2265 GCTCGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTAGTGACCTGAATAATGTCA 2324
         || ||||
Db  8526 GTTGGGACCATCCATAAAAATGACCGAACTCTTCAATCCCTTGCTGACCTGAATAATGTAC 8585

Qy  1903 GCCGTGGCGGCGATCTCCTCATTGACTCTCCAGATCACCTCGAGAAAGTCAAGGTCACT 1962
         |||||||||||||||
Db  1082 GCCCATTGGGATCTTCATTGATTCACTCCCAGAGCACATCCAGGTCTATTAAGCTGTT 1141

Qy  2325 GATTCTCAGCTTATAGGACTTATAGAAACTCCGAAGACTGCAGAGAGGCCCTTTGCTTGG 2384
         |||| |||||| ||| |||||||
Db  8586 GTTTTTCTGCCTACGTACAGCAACAATCAAAATCCGAGAGCTACACAAAAAGCACTATGTTTGG 8645

Qy  1963 TCGAGAGAAATTGCGCCTCTGAAAGAAGAACGAGCCACGTGAGCCACGCCGCTGCTGCGCCA 2022
         |||||||||||||||| |||||||||||||||
Db  1142 CAAAGAAGAATTCTCCCCCATGAAGAAGAGAGTAAGTTTGGTGAATGATTCTTGGCCACCA 1201

Qy  2385 ATCTCTT 2391
         |||||||
Db  8646 ATCTCTT 8652

Qy  2023 GCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCCAGCACTCTGGAAGACCTGAA 2082
         |||||||||||||||||||||||||||||
Db  1202 ACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAATTCCCAGGCCCTGGAACACAGATCAA 1261

Qy  2083 CACCAGATGGAAGCTTCTGCAGGTGGCCGTGCAGGACAGCAGTCATGAAGC 2142
         ||||||| ||  |||||||
Db  1262 CGTCCGATGGAAACAACTACAGGCGCTCAGTTGATGAGAGGCTTCCAGGATGC 1321

Qy  2143 CCACAGGGACTTTGGTCCAGCATCTCCAGCACTTCTTCCACGTCTGTCCAGGGTCCCTG 2202
         ||||  |||||||||||| |||||||||||||
Db  1322 CCACCGGGGACTTTGGGCCTGGGTCACAGCACTTTCCTCCTCGTCCCTGTTCCCTG 1381

Qy  2203 GGAGAGAGCCATCTCGCCAAACAAAGTGCCCAAACAAAGTGCCTACTATATCAACCACCAGACTCAAACAAC 2262
         |||| || ||||||
Db  1382 GGAAAGAGCAGCAATTCACCCAATAAAGTCCCATACATGACCAGAGTTATACCAACCCTAGTCTGACACACAT 1441

Qy  2263 TTGCTGGGACCATCCCAAAATGACAGAGCTCTTTAGCTGACCTGTGACCTTGACTTGAATAATGT 2322
         |||||||||||||| |||||||||||||||| ||||||||
Db  1442 ATGCTGGGACCATCCCAGATGACCAGAGTGTTATACCAAACCCTAGTCTGATCTGACAACACAT 1501

Qy  2323 CAGATTCTCAGCTTATAGGACTTATAGAAACTCCGAAGACTG 2382
         ||||||||||||||| |||||
Db  1502 TAAGTTCTCAGCTTATCCGCCATGAACTCCGCAGAGTCCAGAAAGCCCTGCGCTT 1561

Qy  2383 GGATCTCTTGA 2393
```

Db   1562 GGACCTGGTAA 1572
             ||||| ||

RESULT 8
US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-091-501B-5

Query Match   3.3%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 1.6e-14;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy   865 CCTAAAACGCCAAGTACAACACATAAGGTGCTTCAAGAAGATCTAGAAGAAGAACAAGT 924
             || |||| |||||||||||||||||||||| ||| || ||| ||||||||||| ||
Db   16 CCTGCAAAACCTGCTTGAAGAACATAAAAGTTGCAAAGTTGCAAAGTTGCAAAGCAGGT 75

Qy   925 CAGGGTCAATTCTCTCACTCACATGGTGGTAGTTGATGAATCTAGTGGGGGATCACGC 984
             || ||| ||||||||| || || || || |||| || ||||||||||||| |||
Db   76 GAAGGTGAATTCCTTAACTCATATGTGGTCATTGTGTGGATGAAAACAGTGGGGAGAGCGC 135

Qy   985 AACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGTATTGGGCAAACATCTGTAG 1044
             |||||| || ||| ||| || |||| |||||| ||||||||||||||||||||||
Db   136 CACAGCTGTTTTGGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGCGCGCTGTATGCCG 195

Qy   1045 ATGGA 1049
             |||||
Db   196 CTGGA 200

RESULT 9
US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4

Query Match   3.3%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 2.9e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy   853 TGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCAAGAAGATCTAGA 912
             | |||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   4 TGACCTGCCCTCCCTGCAGAAGCTGCTTCAAGAACATAAAAGTTGCAAAATGACCTTGA 63

Qy   913 ACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGTTGATGAATCTAG 972
             || |||||||| ||||| |||||||||||||||| |||||||||||| ||||| |||
Db   64 AGCTGAACAGTGAAGTAAATTCCTTAACTCACATGGTGGTCATTGTGGATGAAAACAG 123

Qy   973 TGGAGATCACCCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 1032
             ||||||| ||||||| |||| |||||||||||| |||||||||||||||||||||||
Db   124 TGGGGAGTGCCAGAGCTCTTCTGGAAGATCAGTCAGTTACGAAACTGGGTGAGCCTGGAC 183

Qy   1033 AAACATCTGTAGATGGA 1049
             |||||||||||||||||
Db   184 AGCTGTATGCCGTTGGA 200

RESULT 10
US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match   3.3%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 2.9e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy   853 TGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCAAGAAGATCTAGA 912
             | |||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   4 TGATGTGAAATCTTCACAAAAGGTGCTCTAGAGAACATAAAAGTTTGCAAAGTGGATCTTGA 63

Qy   913 ACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTAGTTGATGAATCTAG 972
             || |||||||| ||||| |||||||||||||||| |||||||||||| ||||| |||
Db   64 GGCTGAACAGTGAAGTAAATTCACTAACTCACATGTGTCATGTGGTCATTGTGTGATGAAAACAG 123

Qy   973 TGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 1032
             ||||| ||| ||||| |||| |||||||||||| |||||||||||||||||||||||
Db   124 TGGTGAGAGCCGCTACAGCTATCCCAGCTATCCAGAGAAGACCAGTGTTACAGAAACTTGGTGAGCCGTGGAC 183

Qy   1033 AAACATCTGTAGATGGA 1049
             |||||||||||||||||
Db   184 AGCAGTTATGCCGTTGGA 200

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match   3.2%; Score 76.6; DB 1; Length 7218;
Best Local Similarity   6.7%; Pred.No. 1.5e-12;
Matches   28; Conservative 236; Mismatches 155; Indels 0; Gaps 0;

Qy   513 TTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTCATACTCATGAGGGGTACATGATGG 572
          |||| :: :: ::
Db   1474 TATCTATGCAAGTAGTTAAAGAGATAGAGAATTTGGTACRRRRRRRRRRRRRRRRRRRR 1415

Qy   573 ATTTGACAGCCCATCAGGGCCGGGTTGGTAAATATTCTACAATTGGGAAGCTGATTG 632
          : ::: ::   :: ::: ::
Db   1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355

Qy   633 GAACAGGAAAATTATCAGAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAA 692
          ::: :::::: ::::::: ::::: ::::
Db   1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295

Qy   693 ATTCAAGATGGGAATGCCTCAGGTAGCTAGCGTAGCATGGAAAAACAAAGCAATTTACATAGAG 752
          : ::: ::::: :: ::::: ::: ::::: : :::::
Db   1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235

Qy   753 TTTTAATGGATCTCCAGAATCAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAAACAG 812
          :::: :: ::: : :: ::: :: ::: :::::: ::

Db   1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175

Qy   813 AAGAAAGAACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAAC 872
            :::::::: ::: ::::::::::: :: :::::
Db   1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

Qy   873 GCCAAGTACAACACATAGGTGCTTCAAGAGATCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTC 931
            :: ::: ::: :::::::::: :::::
Db   1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTC 1056

RESULT 12
US-09-687-875A-13
; Sequence 13, Application US/09687875A
; Patent No. 654786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pXX-C2 5' junction
US-09-687-875A-13

Query Match   2.6%; Score 63.6; DB 4; Length 238;
Best Local Similarity 94.3%; Pred. No. 1.4e-09;
Matches   66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1324 TTCCGATGATGCAGTCCTGTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAGTGA 1383
          ||| ||| |||||| ||||||||||||||||| ||||||||||||||||||||||||||
Db   169 TTCGGACGACGCAGTACTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGA 228

Qy   1384 ACTTCGGAAA 1393
          ||||||||||
Db   229 ACTTCGGAAA 238

RESULT 13
US-09-621-976-15639/c
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match   1.9%; Score 46.2; DB 4; Length 505;
Best Local Similarity 17.4%; Pred. No. 0.00059;
Matches   54; Conservative 131; Mismatches 124; Indels 2; Gaps 1;

Qy   619 AAGTAAGCTGATTGGAACAGCAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCA 678
          ||:|||| | ::: : :: | : | :: :::::::::

Db   445 AATYAACRWAWAWGAGAAWKRWKNAGSMRACARAGTTNAWGACWAMARRGWRCARGR--T 388

Qy   679 GATGAATCTCCTAAATTCAGATGGGAAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAG 738
          | : | . | |||: | :::: : ::::: : ::: : ::|: :: |
Db   387 GSWGSWKGGYRWWGRGMAAAKMRRWAAAGSYCGMTSYTSGSKWTGRKGSMTKRKRMWTY 328

Qy   739 CAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACTGAAAGAGTTGAATGACTG 798
          : ::::| : :::::|:|:||:| :::::: :: ::::: :: : :::::|
Db   327 SGMWWTSYKCTKTGKXYTGWKSKKTRWTCTSWRKYKYMWMSGCWARSMKSWARSWYSMAC 268

Qy   799 GCTAACAAAAACAGAAGAAGAACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCT 858
          :::: ||: | :: |: || | ::|| || : | |: |: :|| ||
Db   267 WCMMSASAYRARRSMYGARRSMMRAGAGWWRARGKKRARGKSSMMRSKMSSMRMSAGKA 208

Qy   859 TGAAGACCTAAAACGCCAAGTACACAACAACATAAGCTGCTTCAGAAGATCTAGAACAAGA 918
          :::: |||:: :|: :|: || :: :| | :: ::|:: : :|:|:
Db   207 RMCRRWMWSCRRMSYSCMGSKCMSCRGTCAKMWRYARYAKRYASSMGKYMMGCRWCYAKC 148

Qy   919 ACAAGTCAGGG 929
          : |::||:|||:
Db   147 ARMYGYRSRS 137

RESULT 14
US-09-668-313A-10

; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
US-09-668-313A-10

Query Match   1.8%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.0078;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 2201 TGGGAGGAGAGCCATCTCGCCAAACAAAGTGCCCACTATATCAACCACCGAGACTCAAACA 2260
          ||||| ||| ||||| ||| || ||| | ||: || || |:| |||| | || || ||
Db 1733 TGGGAGGAAGCTTACACAGCAGATGAATCAAGTACTTCATCAACCACGTGACACAGACC 1792

Qy 2261 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 2320
          || |||||| |||| |||| || ||| ||| |||||| | ||| ||| |||| |||
Db 1793 ACGTCCTGGATCCACCCCGTGATGAGCGCCCCTGAACCTGTCCTGTGCAGAGGAGTGAA 1852

Qy 2321 GTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAAG 2361
          |||| || |||| || || ||||| || ||||| ||||
Db 1853 GAGGACTGTCCCAGAGAGCTAACAGACCCGAAAAGCTGATG 1893

RESULT 15
US-09-198-452A-1/c

; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A

; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t

NAME/KEY: misc_feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

Query Match   1.8%;   Score 44;   DB 4;   Length 1230025;
Best Local Similarity 53.5%;   Pred. No. 0.72;
Matches 92;   Conservative   0;   Mismatches   80; Indels   0; Gaps   0;

Qy   642 AATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGAT 701
            ||||| ||| |||| ||||||||| |||| |||| ||||| |||| ||||||| |||
Db   656477 AATTAGAGAAGAAGAGAAGAAGAGAAATTGAGGATATCAAAGACTCAGATACAAAAT 656418

Qy   702 GGGAATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGG 761
            || ||| || ||| || ||| |||| ||| | | ||| | ||| || ||
Db   656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCACTAGGCAAGCAATTTATGTGG 656358

Qy   762 ATCTCCAGAAATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGA 813
            || |||||| || || |||| | ||||
Db   656357 CAATTAAGCAGAAAAACTAAAAGCTTCTAAAGAGACGCGCTGGGAAATAGA 656306

Search completed: April 5, 2004, 08:24:54
Job time : 121.44 secs

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      April  5, 2004, 04:53:30 ; Search time 582.2 Seconds
                               (without alignments)
                               15454.047 Million cell updates/sec

Title:        US-09-845-416-9_COPY_600_3000
Perfect score: 2401
Sequence:     1 acatgcattcaacatcgcca.........tggatctcctgagcctgtca 2401

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched:     270430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters:    4940860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Published Applications NA:*
          1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
          2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
          3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
          4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
          5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
          6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
          7: /cgn2_6/ptodata/2/pubpna/US05_PUBCOMB.seq:*
          8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
          9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
         10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
         11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
         12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
         13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
         14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
         15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
         16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
         17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
         18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	% Score	Query Match	Length	DB	ID	Description
1	2401	100.0	3858	10	US-09-845-416-9	Sequence 9, Appli
2	2401	100.0	4825	10	US-09-845-416-29	Sequence 29, Appl
3	2401	100.0	4848	10	US-09-845-416-35	Sequence 35, Appl
4	2401	100.0	5060	10	US-09-845-416-36	Sequence 36, Appl
5	2250	93.7	3999	10	US-09-845-416-6	Sequence 6, Appli
6	2250	93.7	4966	10	US-09-845-416-28	Sequence 28, Appl
7	2250	93.7	4990	10	US-09-845-416-34	Sequence 34, Appl
8	2067	86.1	4182	10	US-09-845-416-2	Sequence 2, Appl
9	2067	86.1	5149	10	US-09-845-416-27	Sequence 27, Appl
10	1839.4	76.6	5462	15	US-10-149-736-41	Sequence 41, Appl
11	1737	72.3	3531	10	US-09-845-416-10	Sequence 10, Appl
12	1737	72.3	4498	10	US-09-845-416-30	Sequence 30, Appl
13	1725	71.8	5339	15	US-10-149-736-40	Sequence 40, Appl
14	1695	70.6	3510	10	US-09-845-416-12	Sequence 12, Appl
15	1695	70.6	4476	10	US-09-845-416-31	Sequence 31, Appl

16	1419	59.1	5417	16	US-10-149-736-39	Sequence 39, Appl
17	1335.6	55.6	8689	15	US-10-149-736-42	Sequence 42, Appl
18	1335.6	55.6	11058	10	US-09-845-416-1	Sequence 1, Appli
19	1335.6	55.6	11443	15	US-10-149-736-44	Sequence 44, Appl
20	1335.6	55.6	12057	15	US-10-149-736-47	Sequence 47, Appl
21	1335.6	55.6	13957	9	US-09-782-378A-22	Sequence 22, Appl
22	1335.6	55.6	13957	9	US-09-880-107-2284	Sequence 2284, Ap
23	1335.6	55.6	13957	15	US-10-149-736-1	Sequence 1, Appli
24	1335.6	55.6	14069	12	US-10-342-887-434	Sequence 434, App
25	1335.6	55.6	14082	12	US-10-342-887-981	Sequence 981, App
26	1335.6	55.6	14082	12	US-10-341-434-108	Sequence 108, App
27	1323	55.1	2169	10	US-09-845-416-4	Sequence 4, Appli
28	1263	52.6	4414	10	US-09-845-416-32	Sequence 32, Appl
29	1251	52.1	3446	10	US-09-845-416-14	Sequence 14, Appl
30	1143.6	47.6	13815	15	US-10-149-736-2	Sequence 2, Appli
31	1092.2	45.5	1991	10	US-09-845-416-3	Sequence 3, Appli
32	1077	44.9	1667	10	US-09-845-416-7	Sequence 7, Appli
33	976	40.6	1821	10	US-09-845-416-13	Sequence 13, Appl
34	750	31.2	1340	10	US-09-845-416-11	Sequence 11, Appl
35	589	24.5	1434	10	US-09-845-416-15	Sequence 15, Appl
36	495.6	20.6	1096	10	US-10-149-736-4	Sequence 4, Appli
37	487	20.3	10302	9	US-09-782-378A-23	Sequence 23, Appl
38	487	20.3	10302	15	US-10-149-736-3	Sequence 3, Appli
39	475	19.8	16531	14	US-10-101-510-667	Sequence 667, App
40	411.8	17.2	5106	12	US-10-220-120-157	Sequence 157, App
41	387	16.1	387	15	US-10-149-736-32	Sequence 32, Appl
42	348	14.5	348	15	US-10-149-736-31	Sequence 31, Appl
43	331	13.8	333	15	US-10-149-736-9	Sequence 9, Appli
44	327	13.6	327	15	US-10-149-736-8	Sequence 8, Appli
45	324	13.5	324	15	US-10-149-736-33	Sequence 33, Appl

ALIGNMENTS

```
RESULT 1
US-09-845-416-9
; Sequence 9, Application US/0845416
; Publication NO. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
;  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: DE1142
;  CURRENT APPLICATION NUMBER: US/09/845,416
;  CURRENT FILING DATE: 2001-04-30
;  PRIOR APPLICATION NUMBER: 60/200,777
;  PRIOR FILING DATE: 2000-04-28
;  NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 9
;  LENGTH: 3858
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-845-416-9

Query Match       100.0%;  Score 2401;  DB 10;  Length 3858;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy    1 ACATGCATTCAACATCGCCAGATATCATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  600 ACATGCATTCAACATCGCCAGATATCATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 659

Qy   61 TGTTGATACCACCTATCCAGATAAGAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  660 TGTTGATACCACCTATCCAGATAAGAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 719

Qy  121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  720 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779

Qy  181 ACCTAAAGTGACTAAAGAAGAGACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 240
```

```
Db   1860 TCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCCTGGA 1919

Qy   1321 AGGTTCCGATGAGCAGTCCGTTGTTACAAAGACGTTTGGATAACATGACATGACTGGAG 1380
          |||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Db   1920 AGGTTCCGATGATGAGCAGTCCTGTTACAAAGACGTTTGGATAACATGACATGACTGGAG 1979

Qy   1381 TGAACTTCGGAAAAGTCTTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1980 TGAACTTCGGAAAAGTCTTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 2039

Qy   1441 GAAGGCTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2040 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2099

Qy   1501 ATTAAGCCGGCAGCCACCTATTGGAGGGCACTTTCAGAAGCTCAGAAGCAGAACGATGT 1560
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   2100 ATTAAGCCGGCAGCCACCTATTGGAGGGCACTTTCAGAAGCTCAGAAGCAGAACGATGT 2159

Qy   1561 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2160 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2219

Qy   1621 GACTGTACGAATATTCTGACGAGCAGCGCTTTGGAAGGACTAGAGAACTCTACCAGGA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2220 GACTGTACGAATATTCTGACGAGCAGCGCTTTGGAAGGACTAGAGAACTCTACCAGGA 2279

Qy   1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCCTTCTACGAAAGCA 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2280 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCCTTCTACGAAAGCA 2339

Qy   1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAATTGAACCTGCACTCCGCTGACTGGCGAGAG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2340 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCGAGAG 2399

Qy   1801 AAAAATAGATGAGACCCCTGAAAGACTCAAGAGGCCACCGGATGAGCTGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2400 AAAAATAGATGAGACCCCTGAAAGACTCAAGAGGCCACCGGATGAGCTGGA 2459

Qy   1861 CCTCAAGCTGCGCGCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGGCCATCTCCT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460 CCTCAAGCTGCGCGCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGGCCATCTCCT 2519

Qy   1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGAGAAATTGCGCC 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2520 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGAGAAATTGCGCC 2579

Qy   1981 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2580 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2639

Qy   2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACTGAACACCAGATGGAAGCTTCT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2640 TCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACTGAACACCAGATGGAAGCTTCT 2699

Qy   2101 GCAGTGACGTCGTCGAGGACCTCAGGAGACCGAGTCAGGCAGCGAGTGCATGAAGCCCAGGGGACTTTGGTCC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700 GCAGTGACCGTCGAGGACCTCAGGAGACCGAGTCAGGCAGCGAGTGCATGAAGCCCAGGGGACTTTGGTCC 2759

Qy   2161 AGCATCTCAGCACTTTCTTCCACGTCTTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2760 AGCATCTCAGCACTTTCTTCCACGTCTTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2819

Qy   2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGACCATCCCAA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2820 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA 2879

Qy   2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2880 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2939

Qy   2341 GACTGCCATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGGATCTCTTGAGCCGTC 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2940 GACTGCCATGAAACTCCGAAGACTGCGAAAGGCCCTTTGCTTGGGATCTCTTGAGCCCGTC 2999
```

```
Db    780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 839

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 899

Qy    301 GAGCTATGCCTACACACACAGGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 GAGCTATGCCTACACACACAGGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT 959

Qy    361 TCCTTCACAGCATTTGGAAGCTCCGAAGACAAGTCATTTGGCAGTCATTGATGGGAGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    960 TCCTTCACAGCATTTGGAAGCTCCGAAGACAAGTCATTTGGCAGTCATTGATGGGAGAG 1019

Qy    421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1079

Qy    481 TGCTGAGGACACATTGCAAGCAGAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACATTGCAAGCAGAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1139

Qy    541 CCAGTTTCATTACTCATGAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTTCATTACTCATGAGGGGTACATGATTGACAGCCCATCAGGGCCGGGTTGG 1199

Qy    601 TAATATTCTACAATTGGGAAGTAAGCTTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAAGTAAGCTTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA 1259

Qy    661 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 1319

Qy    721 TAGCAATGGAAAAACAAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 TAGCAATGGAAAAACAAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1379

Qy    781 GAAGAGTTGAAATGACTGGCTAACAAAAACAACAGGAAAATGGAGGAAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAGAGTTGAATGACTGGCTAACAAAAACAACAGGAAAATGGAGGAAGA 1439

Qy    841 GCCTCTTGGACCTGATCTTGAAGACCAAGTGCAAGTACAACATCGCCAAGTACTTGCTTCA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 GCCTCTTGGACCTGATCTTGAAGACCAAGTGCAAGTACAACATCGCCAAGTACTTGCTTCA 1499

Qy    901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCAATGGTGGTGTAGT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCATGGTGGTGTAGT 1559

Qy    961 TGATGAATCTAGTGGAGATCACGCAACTTGCTGCTTTGAAGACAACTTAAGGTATTGGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1560 TGATGAATCTAGTGGAGATCACGCAACTTGCTGCTTTGAAGACAACTTAAGGTATTGGG 1619

Qy   1021 AGATCGATGGGCAAAATCATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGACAC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1620 AGATCGATGGGCAAAATCATCTGTAGATGGACAGAGAACCGCTGGGTTCTTTTACAAGACAC 1679

Qy   1081 TCATAGATTACTGCAACAGTTCAACAGTTCTTGCCTGGCTACAGA 1140
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1680 TCATAGATTACTGCAACAGTTCAACAGTTCTTGCCTGGCTACAGA 1739

Qy   1141 AGCTGAAACAACAGCCCAATGCCAATGTCCTACAGGATGCTAAGGAAGGCTCCTAGAAGA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1740 AGCTGAAACAACAGCCCAATGCCAATGTCCTACAGGATGCTAAGGAAGGCTCCTAGAAGA 1799

Qy   1201 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCAAGGTTCCCAAGGTGAAATTGAAGC 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1800 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCAAGGTTCCCAAGGTGAAATTGAAGC 1859

Qy   1261 TCACACAGATGTTTATCACAACCTGGATGAAACAGCCAAAAATCCTGAGATCCCCTGGA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   2401 A 2401
Db   3000 A 3000

RESULT 2
US-09-845-416-29
Sequence 29, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
  APPLICANT: XIAO, XIAO
  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
      THEREOF
  FILE REFERENCE: DE1142
  CURRENT APPLICATION NUMBER: US/09/845,416
  CURRENT FILING DATE: 2001-04-30
  PRIOR APPLICATION NUMBER: 60/200,777
  PRIOR FILING DATE: 2000-04-28
  SOFTWARE: PatentIn Ver. 2.1
  NUMBER OF SEQ ID NOS: 36
SEQ ID NO 29
  LENGTH: 4825
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-845-416-29

Query Match           100.0%;   Score 2401;   DB 10;   Length 4825;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 2401;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
```

```
Qy     1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTGATCCTGAAGA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1357 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTGATCCTGAAGA 1416

Qy    61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1417 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1476

Qy   121 AGTTTTGCTCAACAAGTGACGATTGAGCATTGAAGCCATCCAGGAGTGGAAATGTGAAGC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1477 AGTTTTGCTCAACAAGTGACGATTGAGCATTGAAGCCATCCAGGAGTGGAAATGTGAAGC 1536

Qy   181 ACCTAAAGTGACTAAAGAGAGAACATTTCAGTTACATCAAATGCACTATTCTCAACA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1537 ACCTAAAGTGACTAAAGAGAGAACATTTCAGTTACATCAAATGCACTATTCTCAACA 1596

Qy   241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1597 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1656

Qy   301 GAGCTATGCCTACACACAGGCTGCTGTTATGTCACCACCTGACCCTACACGGAGCCCATT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1657 GAGCTATGCCTACACACAGGCTGCTGTTATGTCACCACCTGACCCTACACGGAGCCCATT 1716

Qy   361 TCCTTCACAGATTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1717 TCCTTCACAGATTTGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1776

Qy   421 TGAAGTAAAACTGGACCGTTATCAAACAGCTTAGAAGAGTATTATCGTGGCTTCTTTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1777 TGAAGTAAAACTGGACCGTTATCAAACAGCTTAGAAGAGTATTATCGTGGCTTCTTTC 1836

Qy   481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy   541 CCAGTTTCATACTCATGAGGGTACATGATGATTGACAGCCCATCAGGGCCCGGTTGG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1897 CCAGTTTCATACTCATGAGGGTACATGATGATTGACAGCCCATCAGGGCCCGGTTGG 1956

Qy   601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1957 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

Qy   661 AACTGAAGTACAAGAGCAGATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2017 AACTGAAGTACAAGAGCAGATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

Qy   721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 2136

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAACAGGAAAAACAACAGGAAAATGAGGAGAAGA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2137 GAAAGAGTTGAATGACTGGCTAACAAAACAGGAAAAACAACAGGAAAATGAGGAGAAGA 2196

Qy   841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 2256

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGCGTGGTGGTAGT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGCGTGGTGGTAGT 2316

Qy   961 TGATGAATCTAGTGGAGACACGCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2317 TGATGAATCTAGTGGAGATCACGCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG 2376

Qy  1021 AGATCGATGGGCAAACATCTGTAGATGGCAACACCCGTGGGTTCTTTTACAAGACAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2377 AGATCGATGGGCAAACATCTGTAGATGGCAACACCCGTGGGTTCTTTTACAAGACAC 2436

Qy  1081 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTCTTGCTGGCTTACAGA 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2437 TCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTCTTGCTGGCTTACAGA 2496

Qy  1141 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGA 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2497 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGCTCCTAGAAGA 2556

Qy  1201 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2557 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAAGGTGAAATTGAAGC 2616

Qy  1261 TCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2617 TCACACAGATGTTTATCACAACCTGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 2676

Qy  1321 AGGTTCCGATGATGCAGTCCTGTTACAACAATGGCAAGACGTTTGGATAACACTCAGTGGAG 1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2677 AGGTTCCGATGATGCAGTCCTGTTACAAGACGTTTGGATAACATGAACTTCAAGTGGAG 2736

Qy  1381 TGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTGGAAGCCAGTTCTGACCAGTG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2737 TGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTGGAAGCCAGTTCTGACCAGTG 2796

Qy  1441 GAAGCGTCGCACCTTTCTCTGCAGGAACTTCGGTGTGGCTACAGCTGAAAGATGATGA 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2797 GAAGCGTCGCACCTTTCTCTGCAGGAACTTCGGTGTGGCTACAGCTGAAAGATGATGA 2856

Qy  1501 ATAAGCCGGCAGGCCACCTATTGGAGGCGACTTTCAGAAGCAGAAGCAGAACGATGT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2857 ATTAAGCCGGCAGGCCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 2916

Qy  1561 ACATAGGGCCTTCAAGAGGGATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2917 ACATAGGGCCTTCAAGAGGGATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2976

Qy  1621 GACTGTACGAATTATTTCTGACAGAGCAGCCTTTGAAGGACTAGAGAAACTTCTACCAGGA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2977 GACTGTACGAATTATTTCTGACAGAGCAGCCTTTGGAGGACTAGAGAAACTTCTACCAGGA 3036

Qy  1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3037 GCCCAGAGAGCTGCCTCCTGAGGAGAGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 3096
```

```
Matches 2401;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy     1 ACATGCATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1380 ACATGCATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 1439

Qy    61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCAATCACATCACTCTTCCA  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1440 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCAATCACATCACTCTTCCA 1499

Qy   121 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGAAGTGGAAATGTTGGCCAAGGCC      180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1500 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGAAGTGGAAATGTTGGCCAAGGCC     1559

Qy   181 ACCTAAAGTGACTAAAGAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA       240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1560 ACCTAAAGTGACTAAAGAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA      1619

Qy   241 GATCACGGTCAGTCTAGCACAGGGTATGCAGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1620 GATCACGGTCAGTCTAGCACAGGGTATGCAGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1679

Qy   301 GAGCTATGCCTACACACAGGCTGCTTGTATGTCACCACCTTCGACCTTACACGGAGCCCATT   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1680 GAGCTATGCCTACACACAGGCTGCTTGTATGTCACCACCTTCGACCTTACACGGAGCCCATT  1739

Qy   361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG    420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1740 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   1799

Qy   421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1800 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  1859

Qy   481 TGCTGAGGACACATTGCAAGCACAAGAGGAGATTTCTAATGATGTGAAGTGGTGAAAGA   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1860 TGCTGAGGACACATTGCAAGCACAAGAGGAGATTTCTAATGATGTGAAGTGGTGAAAGA  1919

Qy   541 CCAGTTCATACTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGGCCGGGTTGG     600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1920 CCAGTTCATACTCATGAGGGTACATGATGGATTGACAGCCCATCAGGGGCCGGGTTGG    1979

Qy   601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1980 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  2039

Qy   661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC    720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2040 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAGATGGGAATGCCTCAGGGTAGC   2099

Qy   721 TAGCATGTGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2100 TAGCATGTGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT  2159

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAACAGGAGAAAATGGAGGAAGA    840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2160 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAACAGGAGAAAATGGAGGAAGA   2219

Qy   841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACAACATAAGGTGCTTCA   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2220 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACAACAACATAAGGTGCTTCA  2279

Qy   901 AGAAGATCTAGAACAAGAACAAGAAGGCTCAGGGTCAATTCTGCTCGTTGGAAGACAACTTAAGGTATTGGG  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2280 AGAAGATCTAGAACAAGAACAAGAAGGCTCAGGGTCAATTCTGCTCGTTGGAAGACAACTTAAGGTATTGGG 2339

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTGCGCTTGCTTGGAAGAACAACTTAAGGTATTGGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2340 TGATGAATCTAGTGGAGATCACGCAACTGCGCTTGCTTGGAAGAACAACTTAAGGTATTGGG 2399

Qy  1021 AGATCGATGGGCAAACATCTGTAGATGGGCAAACATTCTTTTACAAGACAC           1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2400 AGATCGATGGGCAAACATCTGTAGATGGGCAAACATTCTTTTACAAGACAC          2459
```

```
Qy  1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3097 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 3156

Qy  1801 AAAAATAGATAGAGACCCTTGAAAGACTTCAGGAGACTTCAAGAGCCACGGATGAGCTGGA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3157 AAAAATAGATAGAGACCCTTGAAAGACTTCAGGAGACTTCAAGAGCCACGGATGAGCTGGA 3216

Qy  1861 CCTCAAGCGTCGCGCCAAGTGAGGTGATCAAGGATCTGGCAGCCCGTGGGCGGCGATCTCCT 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3217 CCTCAAGCGTCGCGCCAAGTGAGGTGATCAAGGATCTGGCAGCCCGTGGGCGGCGATCTCCT 3276

Qy  1921 CATTGACTCTCTCCAAGATCACCTCGAGAAGGCACTTCGAGAGGAGAAATTGCGCC 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3277 CATTGACTCTCTCCAAGATCACCTCGAGAAGGCACTTCGAGAGGAGAAATTGCGCC 3336

Qy  1981 TCTGAAAGAGAACGTGAGCCACGTCCGCCAGTTACCACTTTGGGCAT 2040
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  3337 TCTGAAAGAGAACGTGAGCCACGTCCGCCAGTTACCACTTTGGGCAT 3396

Qy  2041 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3397 TCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 3456

Qy  2101 GCAGGTGGCCGTCGAGGACCTGCAGGTCGTCATCCGCGG 2160
         |||||||||||||||||||||||||||||||||||||||
Db  3457 GCAGGTGGCCGTCGAGGACCTGCAGGTCGTCATCCGCGC 3516

Qy  2161 AGCAATCTCAGCACTTTCTTTTCTTTCCCACGTCTGTCCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3517 AGCATTCTCAGCACTTTCTTTTCTTTCCCACGTCTGTCCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3576

Qy  2221 AAACAAAGTGCCCTACTATATCAACCACCAGACCAAAACAACTGCTCGGACCATCCCAA 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3577 AAACAAAGTGCCCTACTATATCAACCACCAGACCAAAACAACTGCTCGGACCATCCCAA 3636

Qy  2281 AATGACAGAGAGCTCTACCAGTCTTTAGTCGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3637 AATGACAGAGAGCTCTACCAGTCTTTAGTCGACCTGAATAATGTCAGATTCTCAGCTTATAG 3696

Qy  2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGACCTGTC 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3697 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGACCTGTC 3756

Qy  2401 A 2401
         |
Db  3757 A 3757
```

RESULT 3
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; SOFTWARE: PatentIn Ver. 2.1
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match   100.0%;   Score 2401;   DB 10;   Length 4848;
Best Local Similarity 100.0%;   Pred. No. 0;

```
Qy   2161  AGCATCTCAGCACTTTCTTTCCAGTCTCGTGTCCCTGGGGTCCCTGGGAGAGAGCCATCTCGCC  2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3540  AGCATCTCAGCACTTTCTTTCCAGTCTCGTGTCCCTGGGGTCCCTGGGAGAGAGCCATCTCGCC  3599

Qy   2221  AAACAAAGTGCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA  2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3600  AAACAAAGTGCCTACTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAA  3659

Qy   2281  AATGACAGAGCTCTACCAGTCTTTAGTGACTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG  2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3660  AATGACAGAGCTCTACCAGTCTTTAGTGACTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG  3719

Qy   2341  GACTGCCATGAAACTCCGAAGACTGCAGAAGACTGCAGAAGGCCCTTTGGATCTCTTGAGCCTGTC  2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3720  GACTGCCATGAAACTCCGAAGACTGCAGAAGACTGCAGAAGGCCCTTTGGATCTCTTGAGCCTGTC  3779

Qy   2401  A  2401
           |
Db   3780  A  3780
```

RESULT 4
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

```
Query Match             100.0%;  Score 2401;  DB 10;  Length 5060;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 2401;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1  ACATGCATTCAACATCGCCAGATATCCAGATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1592  ACATGCATTCAACATCGCCAGATATCCAGATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA  1651

Qy     61  TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1652  TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA  1711

Qy    121  AGTTTTGCTCACAAGTGAGTGAGCCATCCAGGAGGTGGAAATGTTGCCAAGGCC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1712  AGTTTTGCTCACAAGTGAGTGAGCCATCCAGGAGGTGGAAATGTTGCCAAGGCC  1771

Qy    181  ACCTAAAGTGACTAAAGAAGAAGAACATTTCAGTTCAGTCACTAATTCCTCAACA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1772  ACCTAAAGTGACTAAAGAAGAAGAACATTTCAGTTCAGTCACTAATTCCTCAACA  1831

Qy    241  GATCACGGTCAGTCTAGCACAGAGGATATGAGCTGGA  300
           |||||||||||||||||||||||||||||||||||||
Db   1832  GATCACGGTCAGTCTAGCACAGAGGATATGAGCTGGA  1891

Qy    301  GAGCTATGCCTACACACAGGCTGCTTATGTCATCGGAGCCCATT  360
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1892  GAGCTATGCCTACACACAGGCTGCTTATGTCATCGGAGCCCATT  1951

Qy    361  TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTGCAGTCATTGATGGAGAG  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1952  TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTGCAGTCATTGATGGAGAG  2011
```

```
Qy   1081  TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460  TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA  2519

Qy   1141  AGCTGAAACAACTGCCAATGTCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2520  AGCTGAAACAACTGCCAATGTCCAATGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA  2579

Qy   1201  CTCCAAGGGAGTAAAAGAGACTGATGAAACAATGGCAGAGACCTCCAAGGTGAAATTGAAGC  1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2580  CTCCAAGGGAGTAAAAGAGACTGATGAAACAATGGCAGAGACCTCCAAGGTGAAATTGAAGC  2639

Qy   1261  TCACACAGATGTTATCACAACAGCCAAAAAATCCTGAGATCCCTGGA  1320
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   2640  TCACACAGATGTTATCACAACAGCCAAAAAATCCTGAGATCCCTGGA  2699

Qy   1321  AGGTTCCGATGATGAGCTCCTGTTACAAAGACGTTTGGATAAACATGAACTTCAAGTGGAG  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700  AGGTTCCGATGATGAGCTCCTGTTACAAAGACGTTTGGATAAACATGAACTTCAAGTGGAG  2759

Qy   1381  TGAACTTCGGAAAAAGTCTCTCAACATTAGTGTCCCATTTGGAAGCCAGTTCTGACCAGTG  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2760  TGAACTTCGGAAAAAGTCTCTCAACATTAGTGTCCCATTTGGAAGCCAGTTCTGACCAGTG  2819

Qy   1441  GAAGCGGTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGTACAGCTGAAAGATGATGA  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2820  GAAGCGGTCTGCACCTTCTCTGCAGGAACTTCTGGTGTGGTACAGCTGAAAGATGATGA  2879

Qy   1501  ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAGCAGAACGATGT  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2880  ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTCCAGCAGTTCAGAAGCAGAACGATGT  2939

Qy   1561  ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2940  ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA  2999

Qy   1621  GACTGTACAGATATTTCTGACAGAGGACAGCGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3000  GACTGTACGAATATTTCTGACAGAGGACAGCGCCCTTTGGAAGGACTAGAGAAACTCTACCAGGA  3059

Qy   1681  GCCCAGAGAGAGCTGCCTCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCTCACGAAAGCA  1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3060  GCCCAGAGAGAGCTGCCTCCTGAGGAGGAGAGCCCAGAATGTCACTCGGCTTCTCACGAAAGCA  3119

Qy   1741  GGCTGAGGAGGTCAAATCTGAGTGGGAAAAATTGAACCTGACCTGCACTCCGCTGACTGGCAGAG  1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3120  GGCTGAGGAGGTCAAATCTGAGTGGGAAAAATTGAACCTGACCTGCACTCCGCTGACTGGCAGAG  3179

Qy   1801  AAAAATAGATGAGACCCCTTGAAAGACTTCAAGAGGCCACGGATGAGCTGGA  1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3180  AAAAATAGATGAGACCCCTTGAAAGACTTCAAGAGGCCACGGATGAGCTGGA  3239

Qy   1861  CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCAGCCCGTGGGCGGATCTCCT  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3240  CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCAGCCCGTGGGCGGATCTCCT  3299

Qy   1921  CATTGACTCTTCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGAAAAAGTTGCGCCC  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3300  CATTGACTCTTCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGAGGAGAAATTGCGCC  3359

Qy   1981  TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGGCCAGCTTACCACTTTGGGCAT  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3360  TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGGCCAGCTTACCACTTTGGGCAT  3419

Qy   2041  TCAGCGTTCACCGTATAACCTTCTCGAAGACCTGGAGAATGGAAGCTTCT  2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3420  TCAGCGTTCACCGTATAACCTTCTCGAAGACCTGAACAACCAGAGGGAAGCTTCT  3479

Qy   2101  GCAGGTGCGGTCGGCCGTCGAGGACCGAGTCAGGCAGCTCAGGCAGGGCCCACAGGGACTTTGGTCC  2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3480  GCAGGTGCGGTCGGCCGTCGAGGACCGAGTCAGGCAGCTCAGGCAGGGCCCACAGGGACTTTGGTCC  3539
```

Qy   421 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGCTTCTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2012 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGCTTCTTTC 2071

Qy   481 TGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2072 TGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 2131

Qy   541 CCAGTTTCATACTACTCCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2132 CCAGTTTCATAC'TCATGAGGGGTACATGATGATTGACAGCCCATCAGGGCCGGGTTGG 2191

Qy   601 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACACAGAAAATTATCAGAAGATGAAGA 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2192 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACACAGAAAATTATCAGAAGATGAAGA 2251

Qy   661 AACTGAAGTACAAGAGCGATGAATCTCTAAATTCAAGATGGGAATGCCTTCAGGGTAGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2252 AACTGAAGTACAAGAGCGATGAATCTCTAAATTCAAGATGGGAATGCCTTCAGGGTAGC 2311

Qy   721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 2371

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAAATGGAGGAAGA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2372 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAAATGGAGGAAGA 2431

Qy   841 GCCTCTTGAGCCTGATCTTGAGCAGGGTCAATTCTCCACTCACACATAAGGTGCTTCA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2432 GCCTCTTGAGCCTGATCTTGAGCAGGGTCAATTCTCCACTCACACATAAGGTGCTTCA 2491

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTGAAATCCTGACCATGGTGGTGGTGAGT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2492 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCCTCACTCCACATGGTGGTGGTGAGT 2551

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2552 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 2611

Qy 1021 AGATCGATGGGCAAACATTCGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2612 AGATCGATGGGCAAACATTCGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAC 2671

Qy 1081 TCATAGATTACTGCCAACATGCCAATGTCCTACAGGATGCTGCCCCTGGACCTGGCTTACAGA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2672 TCATAGATTACTGCCAACATGCCAATGTCCTACAGGATGCTGCCCCTGGAAAGTTCTTGCCTGGCTTACAGA 2731

Qy 1141 AGCTGAAACAACAACTGCCAATGTCCTACCCGTACCCGTAAGGAGAAGGCTTCCTAGGAAGA 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2732 AGCTGAAACAACAACTGCCAATGTCCTACCCGTACCCGTAAGGAGAAGGCTTCCTAGGAAGA 2791

Qy 1201 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGAACCTCCAAGGTGAAATTGAAGC 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2792 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGAACCTCCAAGGTGAAATTGAAGC 2851

Qy 1261 TCACACAGATGTTTATCAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2852 TCACACAGATGTTTATCAACCTGGATGAAAAACAGCCAAAAAATCCTGAGATCCCTGGA 2911

Qy 1321 AGGTTCCGATGATGCAGTGTCCTGTTACAAAGACGTTTCGATAACATGAACTTCAAGTGGAG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2912 AGGTTCCGATGATGCAGTGTCCTGTTACAAAGACGTTTCGATAACATGAACTTCAAGTGGAG 2971

Qy 1381 TGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTGGAAGCCAGTTCTGACCAGTG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2972 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTG 3031

Qy 1441 GAAGCGTCTGCACCTTTCTCTGCGGAACTTCTGGGTCGTGGCTACAGCTGAAAGAATGATGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3032 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGGTCGTGGCTACAGCTGAAAGAATGATGA 3091

Qy 1501 ATTAAGCGCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3092 ATTAAGCGCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 3151

Qy 1561 ACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3152 ACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 3211

Qy 1621 GACTGTACGAATATTCTGACAGAGCAGCCTTGGAAGGACTAGAGAAACTCTACCAGGA 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3212 GACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGA 3271

Qy 1681 GCCCAGAGAGCGCTGCCTCCTGCTGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3272 GCCCAGAGAGCTGCCTCCTGCTGAGAGCCCAGAATGT'CACTCGGCTTCTACGAAAAGCA 3331

Qy 1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3332 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 3391

Qy 1801 AAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGAGGAGTCGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3392 AAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGAGGAGTCGGA 3451

Qy 1861 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGATCCAGCCCGTGGGCAGCGGCCAGCTTCCT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3452 CCTCAAGCTGCGCCAAGCTGAGTGATCAAGGATCCAGCCCGTGGGCGGCCAGCGGCGATCCTCCT 3511

Qy 1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3512 CATTGACTCTCTCCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCAT 3571

Qy 1981 TCTGAAAGAACGTGAGCGTGAGCCGAGGTCAATGAGCACTCTGCTTGCCAGCTTACCACTTGGTGCAT 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3572 TCTGAAAGAACGTGAGCGTGAGCCGAGGTCAATGAGCACTCTGCTTGCCAGCTTACCACTTTGGTGCAT 3631

Qy 2041 TCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACTCTGAACACCAGATGGAAGCTTCT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3632 TCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACTCTGAACACCAGATGGAAGCTTTCT 3691

Qy 2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTCAGGCTCCGATGAGAGCCCACAGGGACTTTGGTCC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3692 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTCAGGCTCCGATGAGAGCCCACAGGGACTTTGGTCC 3751

Qy 2161 AGCATCTCAGCACTTCTTCCACGTCGTGTCCGAGGGTCCCTGGGAGAGCCATCTCGCC 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3752 AGCATCTCAGCACTTCTTCCACGTCGTGTCCGAGGGTCCCTGGGAGAGCCATCTCGCC 3811

Qy 2221 AAACAAAGTGCCCTACTATATCAACCGAGACTCAAACAACTTGCTGGGACCATCCCAA 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3812 AAACAAAGTGCCCTACTATATCAACCGAGACTCAAACAACTTGCTGGGACCATCCCAA 3871

Qy 2281 AATGACAGAGCTGCTCACCAGTCTTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3872 AATGACAGAGCTGCTCACCAGTCTTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3931

Qy 2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCGTGTC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3932 GACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCGTGTC 3991

Qy 2401 A 2401
          |
Db 3992 A 3992

RESULT 5
US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US2003071312A1
; GENERAL INFORMATION:
;   APPLICANT: XIAO, XIAO
;   TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;   TITLE OF INVENTION: THEREOF

```
FILE REFERENCE: DB1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
SOFTWARE: PatentIn Ver. 2.1
NUMBER OF SEQ ID NOS: 36
SEQ ID NO 6
LENGTH: 3999
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-6

Query Match       93.7%;  Score 2250;  DB 10;  Length 3999;
Best Local Similarity   94.5%;  Pred. No. 0;
Matches 2401;  Conservative 0;  Mismatches 0;  Indels 141;  Gaps 1;
```

```
Qy    1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAAACTACTCGATCCTGAAGA    60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAAACTACTCGATCCTGAAGA   659

Qy   61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  660 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA  719

Qy  121 AGTTTTGCCTCAACAGTGAGCATTGAAGCCATCCAGGAGTGAAATGTTGCCAAGGCC    180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  720 AGTTTTGCCTCAACAGTGAGCATTGAAGCCATCCAGGAGTGAAATGTTGCCAAGGCC    779

Qy  181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTACATCATCAAATGCACTATTCTCAACA   240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  780 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTACATCATCAAATGCACTATTCTCAACA   839

Qy  241 GATCACGGTCAGTCTGCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  840 GATCACGGTCAGTCTGCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA  899

Qy  301 GAGCTATGCCTACACACCAGGTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  900 GAGCTATGCCTACACACCAGGTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   959

Qy  361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  1019

Qy  421 TGAAGTAAACCTGGACCGTTATCAAACAGTTTAGAAGAGAGATTTCTAATGATGTGGAAGTATATCGTGGCTTCTTTC  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1020 TGAAGTAAACCTGGACCGTTATCAAACAGTTTAGAAGAGAGATTTCTAATGATGTGGAAGTATATCGTGGCTTCTTTC 1079

Qy  481 TGCTGAGGACACATTGCAAGCAACAACGGAGATTGGAAGCAACAAGGAGATTCTAAATGGGAAAGA  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1080 TGCTGAGGACACATTGCAAGCAACAACGGAGATTGGAAGCAACAAGGAGATTCTAAATGGGAAAGA 1139

Qy  541 CCAGTTTCATCATCTGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCCGGGTTGG   600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1140 CCAGTTTCATCATCTGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCCGGGTTGG  1199

Qy  601 TAATATTCTACAATTGGGAAGTGAAGTAAGCTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1200 TAATATTCTACAATTGGGAAGTGAAGTAAGCTGATTGGAACACAGGAAAATTATCAGAAGATGAAGA 1259

Qy  661 AACTGAAGTACAAGAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1260 AACTGAAGTACAAGAGAGCAGAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  1319

Qy  721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCCCAGAATCAGAAACT    780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGATCCCAGAATCAGAAACT   1379

Qy  781 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAGAACAAGGAAAATGAGGAGA    840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1380 GAAAGAGTTGAATGACTGGCTAACAAAAACAAGAAGAACAAGGAAAATGAGGAGA   1439
```

```
Qy  841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACACATAAGGTGCTTCA    900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1440 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACACATAAGGTGCTTCA   1499

Qy  901 AGAAGATCTAGACAAGAACAAGTCAGGTCAATTCTCCTACTCAGACTGTTACTCT    960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1500 AGAAGATCTAGACAAGAACAAGTCAGGTCAATTCTCCTCACTCAGACTGTTACTCT   1559

Qy  961 TGATGAATCTAGTGGGACATCACGCAACTGCTGCTTTGGAAGAACAACTAAGGTATTGGG   1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1560 TGATGAATCTAGTGGGACATCACGCAACTGCTGCTTTGGAAGAACAACTAAGGTATTGGG   1619

Qy 1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCGTCGGGTTCTTTTACAAGAC--  1078
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1620 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGGGTTCTTTTACAAGACCA   1679

Qy 1079 ----------------------------------------------------------    1078
Db 1680 GCCTGACCTAGCTCCTGACTGACCTATTGGAGCCTTCTCCTACTCAGACTGTTACTCT   1739

Qy 1079 ----------------------------------------------------------    1078
Db 1740 GGTGACACAACCTGTGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTC   1799

Qy 1079 ---------------ACTCATAGATTACTGCACAGTTCCCCTGACCTTGAAAA   1119
Db 1800 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTCCCCTGACCTTGAAAA   1859

Qy 1120 GTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAGTGTCTTACAGGATGTCTCAGAGG   1179
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1860 GTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAAGTGTCTTACAGGATGTCTCAGAGG  1919

Qy 1180 TAAGGAAAGGCCCTAGAAGACTCCAAGGGAGTAAAGACTGATGAAACAATGCAAGA   1239
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1920 TAAGGAAAGGCCCTAGAAGACTCCAAGGGAGTAAAGACTGATGAAACAATGCAAGA   1979

Qy 1240 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAGTTGATGAAAACAGCCA   1299
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1980 CCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCAACAACACCTGGATGAAAACAGCCA   2039

Qy 1300 AAAAATCCTGAGACTCGAGATCCTGGAAGGTTCCGATGATGCCAGTCAGTCCTGTTACAAAGACGTTTGGA   1359
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2040 AAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCCAGTCAGTCCTGTTACAAAGACGTTTGGA   2099

Qy 1360 TAACATGAACTTCAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGGTCCCATT   1419
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2100 TAACATGAACTTCAAGTGGAGTGAACTTCGGAGAAAAGTCTCTCAACATTAGGTCCCATT   2159

Qy 1420 GGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCGCAGGAACCTTCTGGTGTG   1479
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2160 GGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCGCAGGAACTTCTGGTGTG   2219

Qy 1480 GCTACAGCTGAAAGATGATGAATTAAGCCGGCACCCATCATTGGAGGGGCACTTTCCAGC   1539
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2220 GCTACAGCTGAAAGATGATGAATTAAGCCGGCACCCATCATTGGAGGGGCACTTTCCAGC   2279

Qy 1540 AGTTCAGAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACC   1599
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2280 AGTTCAGAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACC   2339

Qy 1600 TGTAATCATGAGTACTCCTTGAGACTGTACAGAGACTGTGTACATGATATTTCTGACAGAGAATTTGGAAG   1659
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2340 TGTAATCATGAGTACTCCTTGAGACTGTACAGAGACTGTGTACATGATATTTCTGACAGAGAATTTGGAAG   2399

Qy 1660 ACTAGAGAAACTCTACCAGGAGCCCCAGAACTGCTCCTGAGGAGAGCGCCCAGAAGT   1719
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2400 ACTAGAGAAACTCTACCAGGAGCCCCAGAACTGCTCCTGAGGAGAGCGCCCAGAAATGT   2459

Qy 1720 CACTCGGCTTCTTCACGAAAGCAGGCTGAGGAGTGGGAAAAATTGAACCT   1779
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2460 CACTCGGCTTCTTCACGAAAGCAGGCTGAGGAGTGGGAAAAATTGAACCT   2519

Qy 1780 GCACTCCGCTGACTGGCGAGAAAATAGATGAGACCCTGAAAGACTTCCAGAACTTCA   1839
```

```
Qy     61 TGTTGATCACCACCTATCCAGATAAGAAGTCATCTTAATGTACATCACACTCTTCCA    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1417 TGTTGATCACCACCTATCCAGATAAGAAGTCATCTTAATGTACATCACACTCTTCCA   1476

Qy    121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATTCCAGGAAGTGGAAATGTTGCCAAGGCC    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATTCCAGGAAGTGGAAATGTTGCCAAGGCC   1536

Qy    181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGACTATTCTCAACA    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1537 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGACTATTCTCAACA   1596

Qy    241 GATCACGGTCAGTCTAGCACACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 GATCACGGTCAGTCTAGCACACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA   1656

Qy    301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   1716

Qy    361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCTTTAGAGAAGAAGTATTATCGTGCTTCTTC    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCTTTAGAGAAGAAGTATTATCGTGCTTCTTC   1776

Qy    421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAGATTTCTAATGATGTGAAGTGGTGAAAGA   1836

Qy    481 TGCTGAGGACACATTGCAAGCACAAGGAGCACAAGGAGATTTCTAATGATGTGCAGGCCATCAGGGGCGGGTTGG    540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGGAGCACAAGGAGATTTCTAATGATGTGCAGGCCATCAGGGGCGGGTTGG   1896

Qy    541 CCAGTTTCATACTCATGAGGGGTACATGAGTGATTGGAACAGGAAAATTATCAGAAGATGAAGA    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTTCATACTCATGAGGGGTACATGAGTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   1956

Qy    601 TAATATTCTACAATTGGGAAGTAAGCTAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGTAAGCTAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   2016

Qy    661 AACTGAAGTACAAGAGCAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC    720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   2076

Qy    721 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAATGAATCTCCAGAATCAGAAACT    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAATGAATCTCCAGAATCAGAAACT   2136

Qy    781 GAAAGAGTTGAATGACTGCTAACAAAAACAGAAGAAATGGAGGAAGA    840
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTTGAATGACTGCTAACAAAAACAGAAGAAATGGAGGAAGA   2196

Qy    841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACACAACATAAGGTGCTTCA    900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGGCCAAGTACACAACATAAGGTGCTTCA   2256

Qy    901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACTTAAGGACAACTTAAGGTATTGGG    960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACTTAAGGACAACTTAAGGTATTGGG   2316

Qy    961 TGATGAATCTAGTGGAATCATCTGTATGTGACATTCGTGATGACGAAGAGACAGAGA    1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TGATGAATCTAGTGGAATCATCTGTATGTGACATTCGTGATGACGAAGAGACAGAGA   2376

Qy   1021 TAGACGAATGACATCTGTAGATGACAAACATCGTGTAGATGACGAAGAGACAGAGACCTGGTTCTTTTACAAGAC--    1078
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 AGATCGATGGGCAAACATCGTGTAGATGACGAAGAGACAGAGACCTGGTTCTTTTACAAGACCA   2436

Qy   1079 -----------------------------------------------------------    1078

Db   2437 GCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTGACTGTTACTCT   2496

Qy   1079 -----------------------------------------------------------    1078
```

```
Db   2520 GCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCA   2579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1840 AGAGGCCACGATGAGTGGACCTGGACCTGCGCGCCCAAGCTGCGCGCCCAAGCTGGGATCCTG   1899

Db   2580 AGAGGCCACGATGAGTGGACCTGGACCTGCGCGCCCAAGCTGCGCGCCCAAGCTGGGATCCTG   2639
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1900 GCAGCCCGTGGGCGATCTCTCATTGACTCTTCCAAGATCACCTCGAGAAAGTCAAGGC   1959

Db   2640 GCAGCCCGTGGGCGATCTCTCATTGACTCTTCCAAGATCACCTCGAGAAAGTCAAGGC   2699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1960 ACTTCGAGAAGAACGTGAGCACGTCGAGAAGAACGTGAGCCACGTCAATGACCTTGCTCG   2019

Db   2700 ACTTCGAGAAGAACGTGAGCACGTCGAGAAGAACGTGAGCCACGTCAATGACCTTGCTCG   2759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2020 CCAGCTTACCACTTTGGGCATTCAGCTCTCTCACCGTATAACCTCAGCACTCTGAAAGACCT   2079

Db   2760 CCAGCTTACCACTTTGGGCATTCAGCTCTCTCACCGTATAACCTCAGCACTCTGAAAGACCT   2819
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2080 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCCAGTCAGGCAGCTGCATGA   2139

Db   2820 GAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCCAGTCAGGCAGCTGCATGA   2879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2140 AGCCCACAGGGACTTGGTCCAGATCTCCAGCACTTCTTCTTCCACGTCTGTCCAGGGTCC   2199

Db   2880 AGCCCACAGGGACTTGGTCCAGATCTCCAGCACTTCTTCTTCCACGTCTGTCCAGGGTCC   2939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2200 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACAGACTCAAAC   2259

Db   2940 CTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACAGACTCAAAC   2999
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2260 AACTTGCTGGGACACCATCCCAAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA   2319

Db   3000 AACTTGCTGGGACACCATCCCAAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAA   3059
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2320 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCCTTG   2379

Db   3060 TGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCCTTG   3119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2380 CTTGGATCTCTTGAGCCTGTCA   2401

Db   3120 CTTGGATCTCTTGAGCCTGTCA   3141
```

RESULT 6
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match   93.7%; Score 2250; DB 10; Length 4966;
Best Local Similarity   94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

```
Qy      1 ACATGCATTCAACATCGCCGAGATATCAATTCAATTAGGCATAGAGAAAACTACTCCGATCCTGAAGA     60
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1357 ACATGCATTCAACATCGCCAGATATCAATTCAATTAGGCATAGAGAAAACTACTCCGATCCTGAAGA   1416
```

```
RESULT 7
US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
;  APPLICANT: XIAO, XIAO
;  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: DE1142
;  CURRENT APPLICATION NUMBER: US/09/845,416
;  CURRENT FILING DATE: 2001-04-30
;  PRIOR APPLICATION NUMBER: 60/200,777
;  PRIOR FILING DATE: 2000-04-28
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
;  LENGTH: 4990
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-845-416-34

  Query Match             93.7%;  Score 2250;  DB 10;  Length 4990;
  Best Local Similarity   94.5%;  Pred. No. 0;
  Matches 2401;  Conservative    0;  Mismatches    0;  Indels  141;  Gaps     1;
```

Qy   1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAAACTACTCGATCCTGAAGA   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAAACTACTCGATCCTGAAGA 1440

Qy   61 TGTTGATACCACCTATCCAGATAAGAGAGTCATCTTAATGTACATCACATCACTCTTCCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TGTTGATACCACCTATCCAGATAAGAGAGTCATCTTAATGTACATCACATCACTCTTCCA 1500

Qy   121 AGTTTTGCCTCAACAAGTGAGCATCCAGCAGCATCCGAAGAAGTGGAAATGTTGCCAAGGCC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AGTTTTGCCTCAACAAGTGAGCATCCAGCAGCATCCGAAGAAGTGGAAATGTTGCCAAGGCC 1560

Qy   181 ACCTAAAGTGACTAAAGAGAACACATTTCAGTTACATCAAATGCACTATTCTCAACA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACCTAAAGTGACTAAAGAGAACACATTTCAGTTACATCAAATGCACTATTCTCAACA 1620

Qy   241 GATCACGGTCAGTCTAGCACAGGCGATATGAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GATCACGGTCAGTCTAGCACAGGCGATATGAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 1680

Qy   301 GAGCTATGCCTACCACACAGGCTGTCTTATGTCACACGGAGCCCATT 360
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GAGCTATGCCTACCACACAGGCTGTCTTATGTCACACGGAGCCCATT 1740

Qy   361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 420

This page contains a nucleotide sequence alignment (query "Qy" vs database "Db") in two columns of blocks, with position numbers along the margins.

RESULT 8
JS-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
GENERAL INFORMATION:
     APPLICANT: XIAO, XIAO
     TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
     TITLE OF INVENTION: THEREOF
     FILE REFERENCE: DE1142
     CURRENT APPLICATION NUMBER: US/09/845,416
     CURRENT FILING DATE: 2001-04-30
     PRIOR APPLICATION NUMBER: 60/200,777
     PRIOR FILING DATE: 2000-04-28
     NUMBER OF SEQ ID NOS: 36
     SOFTWARE: PatentIn Ver. 2.1
     SEQ ID NO 2
     LENGTH: 4182
     TYPE: DNA
     ORGANISM: Homo sapiens
JS-09-845-416-2

Query Match   86.1%; Score 2067; DB 10; Length 4182;
Best Local Similarity 88.1%; Pred. No: 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

Qy   1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   659

Qy   61 TGTTGATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACCATCACTCCTCCA   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660 TGTTGATACCACCTATCCAGATAAGAGTCCATCTTAATGTACATCACCATCACTCCTCCA   719

Qy   121 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   779

Qy   181 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA   839

Qy   241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   840 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA   899

Qy   301 GAGCTATGCCTACCACACAGGCTGCTTATGTCACCACCCTCTGACCCTACACGGAGCCATT   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACCACACAGGCTGCTTATGTCACCACCCTCTGACCCTACACGGAGCCCATT   959

Qy   361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTCATTGATGGAGG   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTCATTGATGGAGG   1019

Qy   421 TGAAGTAAACCTGGACACATTGCAAGCACAGAGCACAGACGCTTTAGAAGAAGCTTTAG   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGGACACATTGCAAGCACAGAGCACAGACGCTTTAGAAGAAGCTTTAG   1079

Qy   481 TGCTGAGGACACATTGCAAGCACAAGGGGTACATGGGGTACATCAGGGGCCGGGGTTGG   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACATTGCAAGCACAAGGGGTACATGGGGTACATCAGGGGCCGGGGTTGG   1139

Qy   541 CCAGTTTCATACTCCATGAGGGGTACATGAGGATTTTGACAGCCCATCAGGGCGTTGG   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTTCATACTCCATGAGGGGTACATGAGGATTTTGACAGCCCATCAGGGCGTTGG   1199

Qy   601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATATCAGAAGATGAAGA   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATATCAGAAGATGAAGA   1259

Qy   661 AACTGAAGTACAAGAGCAGATGAATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAGATGAATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   1319

Qy   721 TAGCATGAAAAACAAAGCAATTACATTTACATAGAGTTTTAATGATCTCCAGAATCAGAAACT   780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 TAGCATGAAAAACAAAGCAATTACATTTACATAGAGTTTTAATGATCTCCAGAATCAGAAACT   1379

Qy   781 GAAAGAGTGAATGACTGCTAACAAAAACAGAGAGAAAACAGGAAGGAAAATGGAGGAAGA   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTGAATGACTGCTGGCTGGTCAACAAAAACAGAGAGAAAACAGGAAGGAAAATGGAGGAAGA   1439

Qy   841 GCCTCTTGGACCTGATCTTGAAGACTAAAACGCCAAGTACAACAACATAAGGTGCTTCA   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCA   1499

Qy   901 AGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCCTCCACTCACGTGGTGGTAGT   960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 AGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATTCCTCCACTCACGTGGTGGTAGT   1559

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG   1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1560 TGATGAATCTAGTGGAGATCACGCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG   1619

Qy   1021 AGATCGATGGGCAAACATCTGTAGATGGACAAGAAGACCGCTGGGTTCTTTCTTTTACAAGAC--   1078
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1620 AGATCGATGGGCAAACATCTGTAGATGGACAAGAAGACCGCTGGGTTCTTTCTTTTACAAGACAT   1679

Qy   1079 --   1078

Db   1680 CCTTCTCAAATGCAAGTGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA   1739

Qy   1079 --   1078

Db   1740 AAAAGAAGATGCAGTGACAACAAGAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT   1799

Qy   1079 --   1078

Db   1800 ATCAAGTCTTCAAAAACTGGCCGATCTAGAAAGAAAAGAAAAGCAATCCAT   1859

Qy   1079 --   1078

Db   1860 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAGTCAGTGAC   1919

Qy   1079 --   1078

Db   1920 CCGAAGACGGAAGCATGGCTCGGATAACTTTGCCCGGTGTTGGGATTAATTTAGTCCAAAA   1979

Qy   1079 ------------ACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA   1116
                        ||||||||||||||||||||||||||||||||||||
Db   1980 ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGA   2039

Qy   1117 AAAGTTTCTTGCCTGGCTTACAGAAGTGAAACAACTGCCAAATGTCCTACAGGATGCTAC   1176
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2040 AAAGTTTCTTGCCTGGCTTACAGAAGTGAAACAACTGCCAAATGTCCTACAGGATGCTAC   2099

Qy   1177 CGGTAAGGAGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA   1236
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2100 CGGTAAGGAGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCA   2159

Qy   1237 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTATCACAACCTGATGAAAACAG   1296
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2160 AGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTATCACAACCTGATGAAAACAG   2219

Qy   1297 CCAAAAATCCTGAGATCCTGAGAATCCTGAGATGCAGTCCTGTTACAAAGACGTTT   1356
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2220 CCAAAAAATCCTGAGATCCTGAGAATGCAGTCCTGTTACAAAGACGTTT   2279

Qy   1357 GGATAACATGAACTTCAGTGAGTGAACTTCGGAAAAGTCTCTCAACATTAGGTCCCA   1416
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2280 GGATAACATGAACTTCAGTGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA   2339

Qy   1417 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTCGAGGAACTTCTGGT   1476
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2340 TTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTCGAGGAACTTCTGGT   2399

```
; TITLE OF INVENTION:  THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE:  2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; SOFTWARE: PatentIn Ver. 2.1
; NUMBER OF SEQ ID NOS: 36
SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27

Query Match          86.1%;    Score 2067;   DB 10;    Length 5149;
Best Local Similarity 88.1%;    Pred. No 0;
Matches 2401;  Conservative 0;  Mismatches 0;  Indels 324;  Gaps    1;

Qy      1 ACATGCATTCAACATCGCCAGATATCAATCATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
            |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1357 ACATGCATTCAACATCGCCAGATATCAATCATTAGGCATAGAGAAACTACTCGATCCTGAAGA 1416

Qy     61 TGTTGATACCACTTATCCAGATAAGAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
            |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   1417 TGTTGATACCACCTATCCAGATAAGAGAGTCCATCTTAATGTACATCACATCACTCTTCCA 1476

Qy    121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 1536

Qy    181 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCAAATGCACTATTCTCAACA 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1537 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCAAATGCACTATTCTCAACA 1596

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGAGACAGTTCATTGGAGTCATTGATGGAGAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 GATCACGGTCAGTCTAGCACAGGGATATGAGACAGTTCATTGGAGTCATTGATGGAGAG 1656

Qy    301 GAGCTATGCCTACACACAGCTGCTTATGTCACCACCTTCACCACCTCGACCGTCACGGAGCCCATT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGCCTACACACAGCTGCTTATGTCACCACCTTCACCACCTCGACCGTCACGGAGCCCATT 1716

Qy    361 TCCTTCACAGCATTGGAAGCTCTGAAGACAAGTCATTGGCAGTTCATTGGAGTCATTGATGATGAGAG 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTGGAAGCTCTGAAGACAAGTCATTGGCAGTTCATTGGAGTCATTGATGATGAGAG 1776

Qy    421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAATTATCGTGGTCTTTTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAGAATTATCGTGGTCTTTTC 1836

Qy    481 TGCTGAGGACACATTGCAAGCACAGGGGTACATGATGGATTTGAACAGCAGAGATTCTAATGGATTGGAAAGA 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAGGGGTACATGATGGATTTGAACAGCAGAGATTCTAATGGATTGGAAAGA 1896

Qy    541 CCAGTTTCATACTCATGAGGGGTACATGGAGGGGTACATGGGCCGGGTTGG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTTCATACTCATGAGGGGTACATGGAGGGGTACATGGGCCGGGTTGG 1956

Qy    601 TAATATTTCTACAATTGGAAGTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTTCTACAATTGGAAGTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

Qy    661 AACTGAAGTACAAGAGCAGATGAATGAATCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 AACTGAAGTACAAGAGCAGATGAATGAATCCTAAATTCAAGATGGAATGCCTCAGGGTAGC 2076

Qy    721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGAGTCTCCAGAATCAGAAACT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGAGTCTCCAGAATCAGAAACT 2136

Qy    781 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAGAAAATGAGGAGGA 840
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGGAGAAAATGAGGAGGA 2196
```

```
Qy   1477 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCACTTTCC 1536
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2400 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGGCACTTTCC 2459

Qy   1537 AGCAGTTCAGAGTCAGAACAGCAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGA 1596
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2460 AGCAGTTCAGAGTCAGAACAGCAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAAAACTAAAGA 2519

Qy   1597 ACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGA 1656
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2520 ACCTGTAATCATGAGTACTTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGA 2579

Qy   1657 AGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAA 1716
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2580 AGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAA 2639

Qy   1717 TGTCACTCCGGCTTCACGAAAGCACAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 1776
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2640 TGTCACTCCGGCTTCACGAAAGCACAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAA 2699

Qy   1777 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 1836
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2700 CCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACT 2759

Qy   1837 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGTGTGATCAAGGGATC 1896
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2760 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGTGTGATCAAGGGATC 2819

Qy   1897 CTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 1956
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2820 CTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAA 2879

Qy   1957 GGCCACTTGAGGAGAAATTGCGCCCTGAAAGAACGTGAGCCACGTCAATGACCTTGC 2016
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2880 GGCACTTGAGGAGAAATTGCGCCCTGAAAGAACGTGAGCCACGTCAATGACCTTGC 2939

Qy   2017 TCGCCAGCTTACCACTTTGGTCTCAGCATTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGA 2076
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2940 TCGCCAGCTTACCACTTTGGTCTCAGCATTCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGA 2999

Qy   2077 CCTGAACACCAGAGATGGAAGCTTCTGCAGGTGGCCGTGCGAGGACCGAGTCAGGCAGCTGCA 2136
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3000 CCTGAACACCAGAGATGGAAGCTTCTGCAGGTGGCCGTGCGAGGACCGAGTCAGGCAGCTGCA 3059

Qy   2137 TGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCCACGTCTGTCCAGGG 2196
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3060 TGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCCACGTCTGTCCAGGG 3119

Qy   2197 TCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTACTATCAACCACGAGACTCA 2256
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3120 TCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTACTATCAACCACGAGACTCA 3179

Qy   2257 AACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAA 2316
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3180 AACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAA 3239

Qy   2317 TAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAACTCCGAGAGGCCCT 2376
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3240 TAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGAACTCCGAGAGGCCCT 3299

Qy   2377 TTGCTTGGATTCTTGAGCCTGTCA 2401
            ||||||||||||||||||||||||
Db   3300 TTGCTTGGATTCTTGAGCCTGTCA 3324
```

```
RESULT 9
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US2003017131 2A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
```

Qy   1597 ACCTGTAATCATGAGTACTCTTGAGACTGTGACTCTTGAGACTGTACGAATATTTCTGACAGCAGCAGCCTTTGGA 1656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3277 ACCTGTAATCATGAGTACTCTTGAGACTGTGACTCTTGAGACTGTACGAATATTTCTGACAGCAGCAGCCTTTGGA 3336

Qy   1657 AGGACTAGAGAAACTCTACCAGGAGCGCTGCCTCCTGAGGAGAGAGCCCAGAA 1716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3337 AGGACTAGAGAAACTCTACCAGGAGCGCTGCCTCCTGAGGAGAGAGCCCAGAA 3396

Qy   1717 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAA 1776
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3397 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAAATTGAA 3456

Qy   1777 CCTGCACTCCGCTACTCGGCAGAGAAAAATAGATGAGACCCTTGAAACTCCAGGGACT 1836
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3457 CCTGCACTCCCGCTGACTCGGCAGAGAGAAAAATAGACCCTTGAAAGACTCCAGGAACT 3516

Qy   1837 TCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 1896
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3517 TCAAGAGGGCCACGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 3576

Qy   1897 CTGGCAGCCCGTGGGCGATCTCCTTCATTGACTCTTCTCCAGATCACCTCGAGAAAGTCAA 1956
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3577 CTGGCAGCCCGTGGGCGATCTCCTTCATTGACTCTTCTCCAGATCACCTCGAGAAAGTCAA 3636

Qy   1957 GGCACTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAGTGACCTTGC 2016
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3637 GGCACTTCGAGGAGAAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3696

Qy   2017 TCGCCAGCTTACCACTTTGGGCATTCAGCTTCGCAGGTGGCCGTCAGGAGTCAGCAGCTGA 2076
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3697 TCGCCAGCTTACCACTTTGGGCATTCAGCTTCGCAGGTGGCCGTCAGGAGTCAGCAGCTGA 3756

Qy   2077 CCTGACACCAGATGGAAGCTTCTGCAGCATCCAGCACTTCTTCTTCCAGTCTGTCCAGGG 2136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3757 CCTGACACCAGATGGAAGCTTCTGCAGCATCCAGCACTTCTTCTTCCAGTCTGTCCAGGG 3816

Qy   2137 TGAAGCCCACAGGACTTTGGTCCAACAGTTGCCCCTGACCTGA 2196
          ||||||||||||||||||||||||||||||||||||||||||
Db   3817 TGAAGCCCACAGGACTTTGGTCCAACAGTTCCCCTGACCTGGG 3876

Qy   2197 TCCCTGGGAGAGAGCCATCTCGCAAACAAAGTGCCCACTACTATCAACCACAGACTCA 2256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3877 TCCCTGGGAGAGAGCCATCTCGCAAACAACAAGTGCCTACTATCAACCACAGACTCA 3936

Qy   2257 AACAACTGCTGGGACCATCCCAAAATGACAGAGCTCACCAGTCTTTAGCTGACCTGAA 2316
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3937 AACAACTGCTGGGACCATCCCAAAATGACAGAGCTCACCAGTCTTTAGCTGACCTGAA 3996

Qy   2317 TAATGTCAGATTCTCAGCTTACAAAACTCGCAGAAACTCGAGAACTCGAGAAGGCCCT 2376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3997 TAATGTCAGATTCTCAGCTTACAAAACTCGCAGAAACTCCGAGAACTCCGAGAGGCCCT 4056

Qy   2377 TTGCTTGGATCTCTTGAGCCTGTCA 2401
          |||||||||||||||||||||||||
Db   4057 TTGCTTGGATCTCTTGAGCCTGTCA 4081

Qy   841 GCCTCTTGGACCTGATCTTGAAGACACCAAGTACAACAACATAAGGTGCTTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 GCCTCTTGGACCTGATCTTGAAGACACCAAGTACAACAACATAAGGTGCTTCA 2256

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGTGGTGGTAGT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGTGGTGGTAGT 2316

Qy   961 TGATGAATCTAGTGGAGATCACCGCAACTCGTGCTTGCTTGGAGAACAACTTAAGGTATTGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 TGATGAATCTAGTGGAGATCACCGCAACTCGTGCTTGCTTGGAGAACAACTTAAGGTATTGGG 2376

Qy   1021 AGATCGATGGGCAAACATCTGTAGAAGGACAGAGAGACCCGTGGGTTCTTTTACAAGAC-- 1078
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 AGATCGATGGGCAAACATCTGTAGAAGGACAGAGAGACCCGTGGGTTCTTTTTACAAGACAT 2436

Qy   1079 --- 1078
Db   ---

Qy   2437 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCCTTTTTGTGCATGGCTTTTCAGA 2496
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 --- 1078

Qy   2497 AAAAGAAGATCGAGTGAACAAGAATTCACACAACTCGCTTTAAAGATCAAAATGAAATGTT 2556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 --- 1078

Qy   2557 ATCAAGTCTTCAAAAACTGGCGCCGTTTAAAAGCCGATCTAGAAAAGAAAGAAATCCAT 2616
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 --- 1078

Qy   2617 GGGCAAACTGTATTCACTCAAACAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2676
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 --- 1078

Qy   2677 CCAGAAGACGGAAGCGATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA 2736
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1079 --------------ACTCATAGATTACTGCAACAGTTCCCCTGACCTGGA 1116

Qy   2737 ACTTGAAAAGAGTACAGCACAGACTTCATAGATTACTGCAACAGTTCCCCTGACCTGA 2796
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2797 AAAGTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAAATGTCCTACAGATGTCTAC 2856

Qy   1117 AAAGTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAAATGTCCTACAGATGTCTAC 1176
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2797 AAAGTTTCTTGCCTGGCTTACAGAGAGCTGAAACAACTGCCAAATGTCCTACAGATGTCTAC 2856

Qy   1177 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCA 1236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2857 CCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCA 2916

Qy   1237 AGACCTTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1296
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2917 AGACCTCCAAGGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 2976

Qy   1297 CCAAAAAATCCTGAGATCCCTGGAAGGTTCCGAGTGAGTGAACTTCAACATTAGGTCCCA 1356
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3037 GGATAACATGAACTTCAAGTGGAGTGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3096

Qy   1357 GGATAACATGAACTTCAAGTGGAGTGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1416
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3037 GGATAACATGAACTTCAAGTGGAGTGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3096

Qy   1417 TTTGGAAGCCAGTGGAAGCCGTCTGACCAGTGGAAGCCTGACCTTCTCTCGCAGGAACTTCTGGT 1476
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3097 TTTGGAAGCCAGTGGAAGCCGTCTGACCAGTGGAAGCCTCGCACCTTCTCTCGCAGGAACTTCTCGT 3156

Qy   1477 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 1536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3157 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCC 3216

Qy   1537 AGCAGTTCAGAAGCAGAACCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 1596
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3217 AGCAGTTCAGAAGCAGAACCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGA 3276

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 5462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-41

Query Match    76.6%;    Score 1839.4;    DB 15;    Length 5462;
Best Local Similarity   88.1%;    Pred. No. 0;
Matches 2115;   Conservative 0;   Mismatches 76;   Indels 210;   Gaps 2;
```

```
Qy      1 ACATGCATTCAACATCGCCAGTATCAAGTATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 ACATGCATTCAACATCGCCAGTATCAAGTATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 858

Qy     61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 918

Qy    121 AGTTTGCCTCAACAAGTGAGCATTGAGCCATTCAGGAAGTGGAAATGTTGCCAAGGCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 AGTTTTGCCTCAACAAGTGAGCATTGAGCCATTCAGGAAGTGGAAATGTTGCCAAGGCC 978

Qy    181 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCATATTCTCAACA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 1038

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGAAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 GATCACGGTCAGTCTAGCACAGGGATATGAAGAACTTCTTCCCCTAAGCCTCGATTCAA 1098

Qy    301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT 1158

Qy    361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGAAGTCATTGGCAGTTCATTGATGGAGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGAAGTCATTGGCAGTTCATTGATGGAGAG 1218

Qy    421 TGAAGTAAACCTGACACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy    481 TGCTGAGGACACATTGCAAGCACAAGGAGACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy    541 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCATCAGGGGCCGGGGTTGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGGCCGGGGTTGG 1398

Qy    601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy    661 AACTGAAGTACAAGAGCAGATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy    721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1558

Qy    781 GAAAGAGTTGAATTGACTGGCTAACAAAAACAGAAGAAACAGAACAAGGAAAATGAGGAGAAGA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1559 ------------------------------------------------------------AC 1560

Qy    841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCA 900
          ||||||||||||||||||||||||||||||||||
Db   1561 CACTATTGGAGCCTCTCCACTCAGCTGTTACTCTGGTGACCACCTGGTTGGGGTTACTAA 1620
```

```
Qy    901 AGAAGATCTAGAACAAGAACAGTCAGGTCAATCTTCCTCACTCACATCATGGTGGTGTAGT 960
Db   1621 GGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTCCTTGATGTTGGAG---------- 1669
Qy    961 TGATGAATCTAGTTGGAGATCACGCAACTGCTTGCTTTGGAAGAAGCAACTTAAGGTATTGGG 1020
Db   1670 ----------------------------------------------------------------- 1669
Qy   1021 AGATCGATTGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGTCTTTTACAAGACAC 1080
Db   1670 ----------------------------------------------------------------- 1669
Qy   1081 TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGA 1140
Db   1670 -CATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCTTACAGA 1728
Qy   1141 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGTCCTAGAAGA 1200
Db   1729 AGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGTCCTAGAAGA 1788
Qy   1201 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGGTGAAAATTGAAGC 1260
Db   1789 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGGCAAGACTTCCAAGGTGAAAATTGAAGC 1848
Qy   1261 TCACACAGATGTTTATCAACAACTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
Db   1849 TCACACAGATGTTTATCAACAACTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1908
Qy   1321 AGGTTCCGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380
Db   1909 AGGTTCCGATGATGCAGTCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAG 1968
Qy   1381 TGAACTTCGGAAAAGTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTG 1440
Db   1969 TGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTG 2028
Qy   1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGTGGCTACAGCTGAAAGATGATGA 1500
Db   2029 -GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTGTGGTGGCTACAGCTGAAAGATGATGA 2088
Qy   1501 ATTAAGCCGGCAGCAGGCGACTTATGGAGGCGACTTCAGAAGCAGTTCAGAAGAGAACGACATGT 1560
Db   2089 ATTAAGCCGGCAGCAGGCGACTTATGGAGGCGACTTCAGCAGTTCAGAAGAGAACGACATGT 2148
Qy   1561 ACATAGGGCTTCAAGAGCGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCCTTGA 1620
Db   2149 ACATAGGGCTTCAAGAGCGAATTGAATCAAAGAACCTGTAATCATGAGTACTCCTTGA 2208
Qy   1621 GACTGTACGAATATTCTGACAGAGCAGCCCTTGGAAGGACTAGAGAAACTCTACCAGGA 1680
Db   2209 GACTGTACGAATATTCTGACAGAGCAGCCCTTGGAAGGACTAGAGAAACTCTACCAGGA 2268
Qy   1681 GCCCAGAGAGCTGCCTCCTGAGGAGGAATGTCACTCGGCTTCACTGAGAAAGCA 1740
Db   2269 GCCCAGAGAGCTGCCTCCTGAGGAGGAATGTCACTCGGCTTCACTGAGAAAGCA 2328
Qy   1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTCGCTGCAGAG 1800
Db   2329 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTCGCTGCAGAG 2388
Qy   1801 AAAAATAGATGAGACCCTTGAAAGTCAGGACTTCAGGAAAGACTTCAGAAGAGGCTGGA 1860
Db   2389 AAAAATAGATGAGACCCTTGAAAGTCAGGACTTCAGGAAAGACTTCAGAAGAGCTGGA 2448
Qy   1861 CCTCAAGCTGCGCAAGTGAGTGATCAAGGATCTGGCAGCCCGTGGGCGATCTCCT 1920
Db   2449 CCTCAAGCTGCGCAAGTGAGTGATCAAGGATCTGGCAGCCCGTGGGCGATCTCCT 2508
Qy   1921 CATTGACTTCTCTCCAGATCACCTCCGAGAAAGTCAAGGCACTTCGAGGACGAAAATTGCGCC 1980
Db   2509 CATTGACTTCTCTCCAGATCACCTCCGAGAAAGTCAAGGCACTTCGAGGACGAAAATTGCGCC 2568
Qy   1981 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2040
```

```
Qy     241 GATCACGGTCAGTCTAGCACAGAGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 GATCACGGTCAGTCTAGCACAGAGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA 899

Qy     301 GAGCTATGCCTACCACACAGGCTGCTTATGTCACCACCTGAACACCAGATGGAAGCTTCT 360
           |||||||||||||||||||||||||||||| |||||||| |||||||||||||||||||||
Db     900 GAGCTATGCCTACCACACAGGCTGCTTATGTCACCACCTGAACACCAGATGGAAGCTTCT 959

Qy     361 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAGTCATTTGGCAGTTCATTGATGAGAG 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     960 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAGTCATTTGGCAGTTCATTGATGGGAGTCC 1019

Qy     421 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTTCTTTC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1020 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGCTTCTTTC 1079

Qy     481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCAATGATGTGGAAGTGGTGAAAGA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCAATGATGTGGAAGTGGTGAAAGA 1139

Qy     541 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1140 CCAGTTTCATACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGG 1199

Qy     601 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACACGAGAAAATTATCAGAAGATGAAGA 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1200 TAATATTCTACAATTGGGAGTAAGCTGATTGGAACACGAGAAAATTATCAGAAGATGAAGA 1259

Qy     661 AACTGAAGTACAAGAGACGATGAATCTTCCAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1260 AACTGAAGTACAAGAGACGATGAATCTTCCAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy     721 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAAACT 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1320 TAGCATGGAAAAACAAGCAATTTACATAGA---------------------------- 1350

Qy     781 GAAAGAGTTGAATGACTGCTAACAAAAACAGAAGAAAGACAAGGAAAATGGAGGAAGA 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1351 ---------------------------------------------------------- 1350

Qy     841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAGGTGCTTCA 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1351 ---------------------------------------------------------- 1350

Qy     901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCCACATGGTGGTGGTAGT 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1351 ---------------------------------------------------------- 1350

Qy     961 TGATGAATCTAGTGGAGATCACGCAACTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1351 ---------------------------------------------------------- 1350

Qy    1021 AGATCGATGGGCAAACATCTGTAGATGGACGAGAGACCGCTCGGGTTCTTTTACAAGACAC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||      ||
Db    1351 ---------------------------------------------------------AC 1352

Qy    1081 TCATAGATTACTGCAACAGTTCCCCCTGACCTGGAAAAGTTTCTTGCCTGCTTACAGA 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1353 TCATAGATTACTGCAACAGTTCCCCCTGACCTGGAAAAGTTTCTTGCCTGCTTACAGA 1412

Qy    1141 AGTTGAAACAACTGCCAATGTCCTACCCGTAAGGAAGGCTTCCTAGAAGA 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1413 AGTTGAAACAACTGCCAATGTCCTACCCGTAAGGAAGGCTTCCTAGAAGA 1472

Qy    1201 CTTCCAAGGAGTAAAGAGCTGATGAAACAATGCCAAGGACCTCCAAGGTGAAATTGAAGC 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1473 CTTCCAAGGAGTAAAGAGCTGATGAAACAATGCCAAGGACCTCCAAGGTGAAATTGAAGC 1532

Qy    1261 TCAACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1533 TCAACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1592

Qy    1321 AGGTTCCGATGATGCAGTCCTGTTACAAGAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380
```

```
Db    2569 TCTGAAAGAGAAGCGTGAGCCACGTCAATGACCTTGCTCGCTGCCAGTCAGTTGGGCAT 2628
Qy    2041 TCAGCTCTCACCGTATAACCTCAGCACTTCTGAAGACTCTGAAGACACCAGATGGAAGCTTCT 2100
Db    2629 TCAGCTCTCACCGTATAACCTCAGCACTTCTGAAGACTCTGAAGACACCAGATGGAAGCTTCT 2688
Qy    2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCGTGCATGAGAGCCCACAGGGACTTTGGTCC 2160
Db    2689 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCGTGCATGAGAGCCCACAGGGACTTTGGTCC 2748
Qy    2161 AGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGGGTCCCTGGGACGGAGAGCCATCTCGCC 2220
Db    2749 AGCATCTCAGCACTTCTTTCCACGTCTGTCCAGGGGTCCCTGGGACGGAGAGCCATCTCGCC 2808
Qy    2221 AAACAAAGTGCCCTACTATATCAACCACCAGACTCAACAACTTGCTGGGACCATCCCAA 2280
Db    2809 AAACAAAGTGCCCTACTATATCAACCACCAGACTCAACAACTTGCTGGGACCATCCCAA 2868
Qy    2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
Db    2869 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2928
Qy    2341 GACTGCCATGAAACTCCGAAGACTGCAGAAGACTGCCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
Db    2929 GACTGCCATGAAACTCCGAAGACTGCCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2988

Qy    2401 A 2401
             |
Db    2989 A 2989
```

RESULT 11
US-09-845-416-10

; Sequence 10, Application US/09845416
; Publication No. US20030173312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10

Query Match   72.3%;   Score 1737;   DB 10;   Length 3531;
Best Local Similarity 86.4%;   Pred. No. 0;
Matches 2074;   Conservative   0;   Mismatches   0;   Indels 327;   Gaps   1;

```
Qy       1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA 659

Qy      61 TGTTGATACCACCACTTCAGATAAGAAGTCCATCTTAATGTACACATCATCACTCTTCCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 TGTTGATACCACCACTTCAGATAAGAAGTCCATCTTAATGTACACATCATCACTCTTCCA 719

Qy     121 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 AGTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 779

Qy     181 ACCTAAAGTGACTAAAGAAGAGAACATTTTCAGTTACATCATCAAATGACTATTCTCAACA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 ACCTAAAGTGACTAAAGAAGAGAACATTTTCAGTTACATCATCAAATGACTATTCTCAACA 839
```

Db   2673 A 2673

RESULT 12
US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; SOFTWARE: PatentIn Ver. 2.1
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 30
;   LENGTH: 4498
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-845-416-30

Query Match   72.3%;   Score 1737;   DB 10;   Length 4498;
Best Local Similarity   86.4%;   Pred No. 0;
Matches 2074;   Conservative   0;   Mismatches   0;   Indels 327;   Gaps   1;

Qy   1 ACATGATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAGAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1357 ACATGATTCAACATCGCCAGATATCAATTCAGATTAGGCATAGAGAAACTACTCGATCCTGAGAG 1416

Qy   61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1417 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA 1476

Qy   121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGCCAAGGCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGCCAAGGCC 1536

Qy   181 ACCTAAAGTGACTAAAGAAGACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1537 ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACTATTCTCAACA 1596

Qy   241 GATCACGGTCAGTCTAGCACAGGGATATGCATCCGGATCCGGACCTGATTCAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 GATCACGGTCAGTCTAGCACAGGGATATGAGAGACTTCTTCCCCTAAGCCTCGATTCAA 1656

Qy   301 GAGCTATGGCTCACACACCAGGTGCTGCTTATGTCACCACCTCTGACCCTACACGGGAGCCCATT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 GAGCTATGGCTCACACACCAGGTGCTGCTTATGTCACCACCTCTGACCCTACACGGGAGCCCATT 1716

Qy   361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1776

Qy   421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1836

Qy   481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1896

Qy   541 CCAGTTTCATCATGAGGGGTACATGATGGATTTGACAGCCATCATGGATGAAGCCGGTTGG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 CCAGTTTCATCATGAGGGGTACATGATGGATTTGACAGCCATCATGGATGAAGCCCGGTTGG 1956

Qy   601 TAATATTCTACAATTGGGAAGTAAGCTGAGTCGATTGGAACAGGAAAATTATCAGAAGATGAAGA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 TAATATTCTACAATTGGGAAGTAAGCTGAGTCGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2016

Qy   661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720

Db   1593 AGGTTCCGATGATGCAGTCTGTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAG 1652
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy

Db   1381 TGAACTTCGGAAAAAGTCTCTCCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1653 TGAACTTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCCAGTTCTGACCAGTG 1712

Qy   1441 GAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGTGTGGCTACAGCTGAAAGATGATGA 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1713 GAAGCGGCTGCACCTTTCTCTGCAGGAACTTCTGTGTGGCTACAGCTGAAAGATGATGA 1772

Qy   1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1773 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAACGATGT 1832

Qy   1561 ACATAGGGCCTTCAAGAGACTAGAGAGCAGCCCTTGGAAGGACTAGAGAAACTCTACCAGGA 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1833 ACATAGGGCCTTCAAGAGACTAGAGAGCAGCCCTTGGAAGGACTAGAGAAACTCTACCAGGA 1892

Qy   1621 GACTGTACGAATATTTCTGACAGCAGACGCCTTTGGAAGAGGAGAAACTTCACCAGGA 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1893 GACTGTACGAATATTTCTGACAGCAGACGCCTTTGGAAGAGCTAGAGAAACTCTACCAGGA 1952

Qy   1681 GCCCAGAGAGCGTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1953 GCCCAGAGAGCGTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAGAAGCA 2012

Qy   1741 GGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2013 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2072

Qy   1801 AAAAATATGAGAGACCCTTGAAAGACTTCCAGGAGCGCACGGATGAGCTGA 1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2073 AAAAATATGAGAGACCCTTGAAAGACTTCCAGGAGCGCACGGATGAGCTGA 2132

Qy   1861 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGATCCTGGCAGCCCGTGGGCGATCTCCT 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2133 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 2192

Qy   1921 CATTGACTCTCTCCAAGATCACCTGAGGCACTTCGAGGAGAGAAATTGCGCC 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2193 CATTGACTCTCTCCAAGATCACCTGAGAGAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 2252

Qy   1981 TCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2253 TCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2312

Qy   2041 TCAGCTTCACGGTATAAACCTCAGCACTCTGGAAGACCTGGAAGACACCAGATGGAAGCTTCT 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2313 TCAGCTTCACGGTATAAACCTCAGCACTCTGGAAGACCTGGAAGACACCAGATGGAAGCTTCT 2372

Qy   2101 GCAGGTGGCCGTCGAGGACCAGGCAGCTGCAGTCAGGGCCCACAGAGCCACTTGGTGGTCC 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2373 GCAGGTGGCCGTCGAGGACCAGGCAGCTGCAGTCAGGGCCCACAGAGCCACTTTGGTGGTCC 2432

Qy   2161 AGCATCTCAGACACTTTCTTTCCAGGTCGTCTGGAGGTCCCTGGGAGAGCCCATCTCGCC 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2433 AGCATCTCAGCACTTTCTTTCCAGGTCGTCTGGAGGTCCCTGGGAGAGCCCATCTCGCC 2492

Qy   2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACCTTGCTGGGACCATCCCAA 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2493 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACCTTGCTGGGACCATCCCAA 2552

Qy   2281 AATGACAGAGCTCTACCAGTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2553 AATGACAGAGCTCTACCAGTCTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2612

Qy   2341 GACTGCCATGAAGAACTCCGAAGACTCCAGAAGGCCCTTGCTTGGATCTCTTGAGCCTGTC 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2613 GACTGCCATGAAGAACTCCGAAGACTCCAGAAGGCCCCTTGCTTGGATCTCTTGAGCCTGTC 2672

Qy   2401 A 2401

```
Db  2770 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCGAGAG 2829
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1801 AAAAATAGATGAGACCCTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2830 AAAAATAGATGAGACCCTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGA 2889

Qy  1861 CCTCAAGTCTGCGCAAGTCGAGTGAGTGATCAAGGATCCTGGCGCCCGTGGCGGATCTCCT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2890 CCTCAAGTCGCGCAAGTCGAGTGAGTGATCAAGGATCCTGGCGCCCGTGGCGGATCTCCT 2949

Qy  1921 CATTGACTTCTCTCCAAGATCACCTCGAGAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2950 CATTGACTTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCC 3009

Qy  1981 TCTGAAAGAGAACGTGAGCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3010 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCAT 3069

Qy  2041 TCAGTCTCACCGTATAACCTCGGAAGACTCACTCTGGAAGACCTGAACACCAGATGAAGCTTCT 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3070 TCAGTCTCACCGTATAACCTCGGAAGACTCACTCTGGAAGACCTGAACACCAGATGAAGCTTCT 3129

Qy  2101 GCAGGTGCCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3130 GCAGGTGGCGTCGAGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCC 3189

Qy  2161 AGCATCTCAGCACACTTCTTTCCACGTCTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3190 AGCATCTCAGCACACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3249

Qy  2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTGCTGGGACCATCCCAA 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3250 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTGCTGGGACCATCCCAA 3309

Qy  2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAACTGAATAATGTCAGATTCTCAGCTTATAG 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3310 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAACTGAATAATGTCAGATTCTCAGCTTATAG 3369

Qy  2341 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTGCTTGGATCTCTTGAGCCTGTC 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3370 GACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTGCTTGGATCTCTTGAGCCTGTC 3429

Qy  2401 A 2401
         |
Db  3430 A 3430
```

RESULT 13
US-10-149-736-40
; Sequence 40, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
;   APPLICANT: Chamberlain, Jeffrey S.
;   APPLICANT: Harper, Scott Q.
;   TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
;   FILE REFERENCE: UM-06968
;   CURRENT APPLICATION NUMBER: US/10/149,736
;   CURRENT FILING DATE: 2002-06-17
;   PRIOR APPLICATION NUMBER: PCT/US01/31126
;   PRIOR FILING DATE: 2001-10-04
;   PRIOR APPLICATION NUMBER: 60/238,848
;   PRIOR FILING DATE: 2000-10-06
;   NUMBER OF SEQ ID NOS: 96
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
;   LENGTH: 5339
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-10-149-736-40

Query Match   71.8%; Score 1725; DB 15; Length 5339;

```
Db  2017 AACTGAAGTACAAGAGCAGCAGATCAATGAATCTTCCTAAATTCAAGATGGGAATGGGAATGCCCAGGGTAGC 2076
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGAATCTCCAGAATCAGAAACT 780
         |||||||||||||||||||||||||||||||||||||||||||||
Db  2077 TAGCATGGAAAAACAAAGCAATTTACATAGA------------------------------- 2107

Qy   781 GAAAGAGTTGAATGACTGCTAACAAAAACAGAGAAGAAGAACAAGGAAAATGGAGGAAGA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2108 ------------------------------------------------------------ 2107

Qy   841 GCCTCTTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2108 ------------------------------------------------------------ 2107

Qy   901 AGAAGAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGGTAGT 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2108 ------------------------------------------------------------ 2107

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2108 ------------------------------------------------------------ 2107

Qy  1021 AGATCGATGGGCAAACATCTGTAGATGAGACAGAAGACACGCTGGGTTCTTTTACAAGACAC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||-AC
Db  2108 --------------------------------------------------AC 2109

Qy  1081 TCATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTGCTTACAGA 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2110 TCATAGATTACTGCAACAGTTCCCCTGACCTGGAAAAGTTTCTTGCCTGCTTACAGA 2169

Qy  1141 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2170 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGA 2229

Qy  1201 CTCCAAGGAGGTAAAGAGCTGATGAAAACAATGGCAAGGACCTCCAAGGTGAAATTGAAGC 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2230 CTCCAAGGAGGTAAAGAGCTGATGAAAACAATGGCAAGGACCTCCAAGGTGAAATTGAAGC 2289

Qy  1261 TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGGA 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2290 TCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAAATCCTGAGATCCCTGGA 2349

Qy  1321 AGGTTCCGATGATGCAGTCCTGTTTACAAAGACGTTTGATAACATGAACTTCAAGTGGAG 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2350 AGGTTCCGATGATGCAGTCCTGTTTACAAAGACGTTTGATAACATGAACTTCAAGTGGAG 2409

Qy  1381 TGAACTTCGGAAAAAGTCTCTCAACACATTAGGTCCCATTGGAGACCCAGTTCTGACCAGTG 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2410 TGAACTTCGGAAAAAGTCTCTCAACACATTAGGTCCCATTGGAGACCCAGTTCTGACCAGTG 2469

Qy  1441 GAAGCGTCTGCACCTTTCCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2470 GAAGCGTCTGCACCTTTCCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2529

Qy  1501 ATTAAGCCGGCAGGCACCTATTGGAGGGACACTTCCAGCAGTTCAGAGAGCAGAACGATGT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2530 ATTAAGCCGGCAGGCACCTATTGGAGGGACACTTCCAGCAGTTCAGAGAGCAGAACGATGT 2589

Qy  1561 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2590 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2649

Qy  1621 GACTGTACGAATATTTCTGACAGAGCAGCGTTTGAAGACTAGAGAAACTCTACCAGGA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2650 GACTGTACGAATATTTCTGACAGAGCAGCGTTTGAAGACTAGAGAAACTCTACCAGGA 2709

Qy  1681 GCCCAGAGAGCTGCCTCCTGAGGAGAGGCCCAGAAGTCACTCGGCTTCTACGAAAGCA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2710 GCCCAGAGAGCTGCCTCCTGAGGAGAGGCCCAGAAATGTCACTCGGCTTCTACGAAAGCA 2769

Qy  1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCGAGAG 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Best Local Similarity 86.1%;  Pred. No. 0;
Matches 2068;  Conservative 0;  Mismatches 0;  Indels 333;  Gaps 1;

Qy      1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA  858

Qy     61 TGTTGATACCCACCTATCCAGATAAGAGAGTCCATCTTAATGTACATCACATCACTCTTCCA  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 TGTTGATACCCACCTATCCAGATAAGAGAGTCCATCTTAATGTACATCACATCACTCTTCCA  918

Qy    121 AGTTTTGCCTCAACAAGTGACCATTGAAGCAGTGAAGTGAAATGTTGCCAAGGCC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 AGTTTTGCCTCAACAAGTGAGCCATTGAAGCAGTGAAGTGAAATGTTGCCAAGGCC  978

Qy    181 ACCTAAAGTGACTAAAGAAGACACATTTCAGTTACTCACTCATCATCAAATGCTACTATTCTCAACA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 ACCTAAAGTGACTAAAGAAGACAACATTTCAGTTACTCACTCATCATCAAATGCACTATTCTCAACA 1038

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAA  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCCTAGATTCAA 1098

Qy    301 GAGCTATGCCTCACACCAGGCTGCTTATGTCAAGACAAGTCACCACGGCCACGGCCCATT  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTCACACCAGGCTGCTTATGTCAAGACAAGTCACCACGGCCACGGCCCATT 1158

Qy    361 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG 1218

Qy    421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTTATCGTTGGCTTCTTTC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTTATCGTTGGCTTCTTTC 1278

Qy    481 TGCTGAGGACACCATTGCAAGCACAAGGACACAAGGGAGATTCTAATGATGTGGAAGTGGTGAAAGA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACCATTGCAAGCACAAGGACACAAGGGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy    541 CCAGTTCATATCTCATGAGGGGTACATGATGGGTACAGCGCCATCAGGGCCGGGTTGG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTCATATCTCATGAGGGGTACATGATGGGTACAGCGCCATCAGGGCCGGGTTGG 1398

Qy    601 TAATATTCTACAAATTGGGAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAAATTGGGAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1458

Qy    661 AACTGAAGTACAAGAGCAGATGAATCAATTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCAATTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT  780

Qy    721 TAGCATGGGAAAACAAAGCAATTTACATAGAGTTTTACAATGGAATCTCCAGAATCAGAAACT  780
          ||||||||||||||||||||||||||||||
Db   1519 TAGCATGGGAAAACAAAGCAATTTACA---------------------------------  1545

Qy    781 GAAAGAGTTGAATGACTGGCTCAACAAAAACAGAAGAAAGAAAATGGAGGAAGA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 ----------------------------------------------------  1545

Qy    841 GCCTCTTGGACCTGATCTTGAAGACTTGAAGAAACCAAGAACAAAACATAAGGTGCTTCA  900
Db   1546 -----------------------------------------------------  1545

Qy    901 AGAAGATCTAGAACAGAGACAAGCAAGTCAGGGTCAATTCTCTTCACTCCACATGGTGGTGGTAGT  960
Db   1546 -----------------------------------------------------  1545

Qy    961 TGATGAATCTAGTGGAGATCACGCAACTGCTTGTTTGGAAGAACAACTTAAGGTTATTGGG 1020
Db   1546 -----------------------------------------------------  1545

Qy   1021 AGATCGATGGGCAACATCTGTAGATGGACAGAAGACACCGCTGGGTTCTTTTTACAAGACAC 1080
Db   1546 -----------------------------------------------------  1545
```

```
Db   1546 -------------------------------------------------------------  1545
Qy   1081 TCATAGATTACTGCAACAGTTCCCCTGACCTTGGCAAAGTTTCTTGCCTGGCTTACAGA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 TCATAGATTACTGCAACAGTTCCCCTGACCTTGGCAAAGTTTCTTGCCTGGCTTACAGA  1605

Qy   1141 AGCTGAAACAACTGCCAATCTCCTACAGATGATGCCAATGTCCTACAGGATGTCCTACAGAAGGCTCCTAGAAGA 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 AGCTGAAACAACTGCCAATGTCCTACAGATGATGCCAATGTCCTACAGGATGTCCTACAGAAGA  1665

Qy   1201 CTCCAAGGAGTAAAGAGCTGATGAAACAATGCAAGACCTCCAAGTGAAATTGAAGC 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 CTCCAAGGGAGTAAAGAGCTGATGAAACAATGCAAGGGTGAAATTGAAGC  1725

Qy   1261 TCACACAGATGTTATCCAACCTGATGAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1726 TCACACAGATGTTATCCAACCTGATGAAACAGCCAAAAAATCCTGAGATCCCTGGA  1785

Qy   1321 AGGTTCCGATGATGGCAGTCCTGTTACAAGACGTTGGATAACATGAACTTCAAGTGGAG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1786 AGGTTCCGATGATGGCAGTCCTGTTACAAGACGTTTGGATAACATGAACTTCAAGTGGAG  1845

Qy   1381 TGAACTTCGGAAAAAGTCTTCTCAACATTAGGTCCCATTGGAGCCAGTTCTGACCAGTG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1846 TGAACTTCGGAAAAAGTCTTCTCAACATTAGGTCCCATTGGAGCCAGTTCTGACCAGTG  1905

Qy   1441 GAAGGTCTGCACCTTCTCTGCAGGAACTCTCGGTGTGGCTACAGCTGAAAGATGATGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1906 GAAGGTCTGCACCTTCTCTGCAGGAACTCTCGGTGTGGCTACAGCTGAAAGATGATGA  1965

Qy   1501 ATTAAGCCGGCAGCCACCTATTGGAGGGCGACTTTCAGCAGTTCAGGAGCAGAACGATGT 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1966 ATTAAGCCGGCAGCCACCTATTGGAGGGCGACTTTCAGCAGTTCAGGAGCAGAACGATGT  2025

Qy   1561 ACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2026 ACATAGGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA  2085

Qy   1621 GACTGTACGAATATTTCTGACAGCAGCTGCCTTGATCAAGAGCAGCCCGTGGGCGGATCTCCT 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2086 GACTGTACGAATATTTCTGACAGCAGCTGCCTTGATCAAGAGCAGCCCGTGGGCGGATCTCCT  2145

Qy   1681 GCCCAGAGAGCTGCCTCCTCGAGGAGAGCCCAGAAATGTCACTCGGCTTCTACGGAAAGCA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2146 GCCCAGAGAGCTGCCTCCTCGAGGAGAGCCCGAAATGTCACTCGGCTTCTACGGAAAGCA  2205

Qy   1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTTCAAGAGACTTCAGGGTAGC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2206 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTTCAAGAGACTTCAGGGTAGC  2265

Qy   1801 AAAAATAGATGAGACCCCTTGAAAGACTTCCAGGAACTTTAATTGGATCTCCAGAATCAGAAACT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2266 AAAAATAGATGAGACCCCTTGAAAGACTTCCAGGAACTTCAAGAGGCCACGGGATCAGAGCTGGA  2325

Qy   1861 CCTCAAGCTGCGCGCCAAGCTGAGGTGATCAAGGCGCACCGTGGGCGCAGCCCGTGGGCGGATCCT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2326 CCTCAAGCTGCGCGCCAAGCTGAGGTGATCAAGGCGCACTCCTGGCAGCCCGTGGGCGGATCTCCT  2385

Qy   1921 CATTGACTCTCTCCAAGATCTCCCAAGAAGTCAAGGCACTTCGAGGAGAATTGCGCC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2386 CATTGACTCTCTCCAAGATCTCCCAAGAAGTCAAGGCACTTCGAGAGAAATTGCGCC  2445

Qy   1981 TCTGAAAGAGAACGTTGAGCCGTATAACCTCGGAGACCTCAATGACGCTTGTGCTTGGGCAT 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2446 TCTGAAAGAGAACGTTGAGCCGTATAACCTCGGAGACCTTGAACACCGGAAGCTTCT  2505

Qy   2041 TCAGCTCTCACCCGTAACCTCAGCAACTCCTGCAGTCAGGCAGGTCATGAAGCCCACAGGGACTTTGGTCC 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2506 TCAGCTCTCACCCGTATAACCTCAGCAACTCCTGCAGTCAGGCAGGTCATGAAGCCCACAGGGACTTTGGTCC  2565

Qy   2101 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTCCATGAAGCCCACAGGGACTTTGGTCC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2566 GCAGGTGGCCGTCGAGGACCGAGTCAGGCAGTCCATGAAGCCCACAGGGACTTTGGTCC  2625
```

```
Db   960  TCCTTCACGCATTTGGAAGCTCCTGAAGACAAGTCATTCGGCAGTTCATTGATGGAGAG   1019

Qy   421  TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAGAGAAGTATTATCGTGGCTTCTTTC   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2626  AGCATCTCAGCACTTCTTCTTCTTTCCACGTCTGTCCAGGGTCCCAGGGAGAGAGCGCC   2685

Qy  2221  AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACAACTGCTGGGACCATCCCAA   2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2686  AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACAACTGCTGGGACCATCCCAA   2745

Qy  2281  AATGACAGAGCTCTTACCAGTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTATAG   2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2746  AATGACAGAGCTCTTACCAGTCTTTAGTGACCTGAATAATGTCAGATTCTCAGCTTATAG   2805

Qy  2341  GACTGCCATGAAACTCCGAAGACTGCGAGAAGGCGCCCTTTGCTTGGATCTCTTGAGCCTGTC   2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2806  GACTGCCATGAAACTCCGAAGACTGCGAGAAGGCGCCCTTTGCTTGGATCTCTTGAGCCTGTC   2865

Qy  2401  A   2401
          |
Db  2866  A   2866
```

RESULT 14
US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030173121A1
; GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match   70.6%;   Score 1695; DB 10; Length 3510;
Best Local Similarity   85.5%;   Pred. No. 0;
Matches 2053; Conservative   0; Mismatches   0; Indels 348; Gaps   1;

```
Qy     1  ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600  ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   659

Qy    61  TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660  TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCA   719

Qy   121  AGTTTGCCTCAACAAGTGAAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720  AGTTTGCCTCAACAAGTGAAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   779

Qy   181  ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCACATTACTCTCAACA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780  ACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCACATTACTCTCAACA   839

Qy   241  GATCACGGTCAGTCTAGCACAGGGGATATGAGACACAGGAACTTCTCCCCGAAGCCTCGATTCAA   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   840  GATCACGGTCAGTCTAGCACAGGGATATGAGACACAGGAACTTCTTCCCCCTTAAGCCTCGATTCAA   899

Qy   301  GAGCTATGCCTACCACAGGCCTGCTTATGTCACCACCTTCACGGAGGCCATT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900  GAGCTATGCCTACCACAGGCCTGCTTATGTCACCACCTTCACGGAGGCCATT   959

Qy   361  TCCTTCACGCATTTGGAAGCTCCTGAAGACAAGTCATTCGGCAGTTCATTGATGGAGAG   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1692  GAAGCGTCGCACCTTTCTCTGCAGAACTTCTGCAGCGTCAGCGAACTGAAAGATGATGA   1751
```

```
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: DE1142
;  CURRENT APPLICATION NUMBER: US/09/845,416
;  CURRENT FILING DATE: 2001-04-30
;  PRIOR APPLICATION NUMBER: 60/200,777
;  PRIOR FILING DATE: 2000-04-28
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: Patentin Ver. 2.1
;  SEQ ID NO 31
;  LENGTH: 4476
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-845-416-31

Query Match           70.6%;  Score 1695;  DB 10;  Length 4476;
Best Local Similarity 85.5%;  Pred. No. 0;
Matches 2053;  Conservative    0;  Mismatches   0;  Indels 348;  Gaps  1;
```

```
Qy      1 ACATGCATTCAACATCGCCAGATATCAATCAATTAGGCATAGGAAACTACTCGATCCTGAAGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1356 ACATGCATTCAACATCGCCAGATATCAATCAATTAGGCATAGGAAACTACTCGATCCTGAAGA 1415

Qy     61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCATCACTCTTCCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1416 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCATCACTCTTCCA 1475

Qy    121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATTCCAGGAAGTGGAAATGTTGCCAAGGCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1476 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATTCCAGGAAGTGGAAATGTTGCCAAGGCC 1535

Qy    181 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCAAATGACTATTCTCAACA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1536 ACCTAAAGTGACTAAAGAAGAACATTTCAGTTACATCAAATGACTATTCTCAACA 1595

Qy    241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTGATTCAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1596 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTGATTCAA 1655

Qy    301 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1656 GAGCTATGCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1715

Qy    361 TCCTTCACAGCATTTGGAAGCTCCTGAGACAAGCTCATTGGACAAGTCATTGATGAGAG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1716 TCCTTCACAGCATTTGGAAGCTCCTGAGACAAGCTCATTGGACAAGTCATTGATGAGAG 1775

Qy    421 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1776 TGAAGTAAACCTGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 1835

Qy    481 TGCTGAGGACACATTGCAAGCACAAGGGGTACATGATGGGTACATGATGGGTGGAAAGA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1836 TGCTGAGGACACATTGCAAGCACAAGGGGTACATGATGGGTACATGATGGGTGGAAAGA 1895

Qy    541 CCAGTTTCATATCTCATGAGGGGTACAGCGCCATCAGGGGCGGGGTTGG 600
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1896 CCAGTTTCATATCTCATGAGGGGTACAGCGCCATCAGGGGCGGGGTTGG 1955

Qy    601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1956 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 2015

Qy    661 AACTGAAGTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2016 AACTGAAGTACAAGAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2075

Qy    721 TAGCATGGAAAAACAAGACAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2076 TAGCATGGAAAAACAAGACAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 2135

Qy    781 GAAAGAGTTGAATGACTGCTAACAAAAAACAGAAGAAAAATGGAGGAAGA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2136 GAAAGAGTTGAATGACTGCTAACAAAAAACAGAAGAAAAATGGAGGAAGA 2195
```

```
Qy   1501 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAGCAGAACGATGT 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1752 ATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAGCAGAACGATGT 1811

Qy   1561 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1812 ACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1871

Qy   1621 GACTGTACGAATATTTCTGACAGAGCAGCCTTGGAAGACTAGAGAAACTCTACCAGGA 1680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1872 GACTGTACGAATATTTCTGACAGAGCAGCCTTGGAAGACTAGAGAAACTCTACCAGGA 1931

Qy   1681 GCCCAGAGAGCAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1932 GCCCAGAGAGCAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTCGGCTTCTACGAAAGCA 1991

Qy   1741 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1800
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1992 GGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2051

Qy   1801 AAAAATAGATGAGACCCTTGAAAGACTTCAAGGAGGCCACGGATGGAGCTGGA 1860
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2052 AAAAATAGATGAGAACGTGAGCCCGTGAGGAGCTGCAGGCCAGTCCAGGAACTTGAA 2111

Qy   1861 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGA 1920
            ||||||||||||||||||||||||||||||||||
Db   2112 CCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGA 2171

Qy   1921 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGAA 1980
            ||||||||||||||||||||||||||||||||||||||||
Db   2172 CATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGAA 2231

Qy   1981 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGCGCCAGCTTACCACTTTGGGCAT 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2232 TCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTGCGCCAGCTTACCACTTTGGGCAT 2291

Qy   2041 TCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2292 TCAGCTCTCACCGTATAAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGAAGCTTCT 2351

Qy   2101 GCAGGTGGCCGCTGCTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCC 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2352 GCAGGTGGCCGCTGCTGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCC 2411

Qy   2161 AGCATCTCAGCACTTTCTTTCCACGTCTGTCGCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2412 AGCATCTCAGCACTTTCTTTCCACGTCTGTCGCCAGGGTCCCTGGGAGAGAGCCATCTCGCC 2471

Qy   2221 AAACAAAAGTGCCCTACTATCAACCAGAGACTCAAACAACTTGCTGGGAACCATCCCAA 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2472 AAACAAAAGTGCCCTACTATCAACCAGAGACTCAAACAACTTGCTGGGAACCATCCCAA 2531

Qy   2281 AATGACAGCAGACTCTTAGCTGACCAGTTGACTTGCTTGGGGTTCAGATTCTCAGCTTATAG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2532 AATGACAGCAGACTCTTAGCTGACCAGTTGACTTGCTTGGGGTTCAGATTCTCAGCTTATAG 2591

Qy   2341 GACTGCCATGAAAACTCCGAGAACTCGAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2400
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2592 GACTGCCATGAAAACTCCGAGAACTCGAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 2651

Qy   2401 A 2401
            |
Db   2652 A 2652
```

```
RESULT 15
US-09-845-416-31
;  Sequence 31, Application US/09845416
;  Publication No. US20030171312A1
;  GENERAL INFORMATION:
;  APPLICANT: XIAO, XIAO
;  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
```

Qy 1921 CATTGACTCTCTCCAAGATCACCTCGAGGAGAAGTCAGGCACTTCGAGGAGAAATTGCGCC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2928 CATTGACTCTCTCCAAGATCACCTCGAGGAGAAGTCAGGCACTTCGAGGAGAAATTGCGCC 2987

Qy 1981 TCTGAAAGAGAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTGGGCAT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2988 TCTGAAAGAGAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCACTTGGGCAT 3047

Qy 2041 TCAGCTCTCACCGTATAACCTCGGAAGACCTGAACACCAGATGGAAGCTTCT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3048 TCAGCTCTCACCGTATAACCTCGGAAGACCTGAACACCAGATGGAAGCTTCT 3107

Qy 2101 GCAGTGCGGCCGTCGAGGACCGAGTCAGGCAGTCGCATGAAGCCACAGGGACTTTGGTCC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3108 GCAGGTGGCGGCCGTCGAGGACCGAGTCAGGCAGTCGCATGAAGCCACAGGGACTTTGGTCC 3167

Qy 2161 AGCATCTCAGCACTTCTTTCCACGTCTGTCCCTGGAGAGCCATCTCGCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3168 AGCATCTCAGCACTTCTTTCCACGTCTGTCCCTGGAGAGCCATCTCGCC 3227

Qy 2221 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTCGGGACCATCCCAA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3228 AAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTCGGGACCATCCCAA 3287

Qy 2281 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3288 AATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAG 3347

Qy 2341 GACTGCCATGAAACTCCGAGACTCGCAGAGGCCCTTGCTTGGATCTTCTTGAGCCTGTC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3348 GACTGCCATGAAACTCCGAGAACTCGCAGAGGCCCTTGCTTGGATCTTCTTGAGCCTGTC 3407

Qy 2401 A 2401
Db 3408 A 3408

Search completed: April 5, 2004, 16:33:42
Job time : 622.2 secs

Qy   841 GCCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACATAAGGTGCTTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2196 GCCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACATAAGGTGCTTCA 2255

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGGTAGT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2256 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGTGGTAGT 2315

Qy   961 TGATGAATTCAGTGGAGATCACGCAACTGCTCGTCTTGGAAGAACAACTTAAGGTATTGGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2316 TGATGAATTCAGTGGAGATCACGCAACTGCTCGTCTTGGAAGAACAACTTAAGGTATTGGG 2375

Qy 1021 AGATGATGGGCAAACATCTGTAGATGACAAGAAGACCGCTGGGTTCTTTTACAAGACAC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2376 AGATCGATGGGCAAACATCTGTAGATGACAAGAAGACCGCTGGGTTCTTTTACAAGA-- 2432

Qy 1081 TCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCTCGGCTTACAGA 1140

Db 2433 --- 2432

Qy 1141 AGCTGAAACAACTGCCAATGTCCTACAGATGTCGTACCCGTAAGGAAAGGCTCCTAGAAGA 1200

Db 2433 --- 2432

Qy 1201 CTCCAAGGGAGTAAAAGAGCTGATGAAACAATGCAAGACCTTCAAGCTGAAATTGAAGC 1260

Db 2433 --- 2432

Qy 1261 TCACACAGATGTTTATCAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGA 1320

Db 2433 --- 2432

Qy 1321 AGGTTCCGATGATGCAGTCCTGTTACAAGAGACGTTTGGATAACATGAACTTCAAGTGGAG 1380

Db 2433 --- 2432

Qy 1381 TGAACTTCCGGAAAAAGTCTCTCAACATTAGGTCCCATTGGAAGCCAGTTCTGACCAGTG 1440
                                              ||||||||||||||
Db 2433 -----------------------------------CAGTTCTGACCAGTG 2447

Qy 1441 GAAGCGTCTGCACCTTTCTCTGCGAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2448 GAAGCGTCTGCACCTTTCTCTGCGAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGA 2507

Qy 1501 ATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCCAGCAGTTCAGAGAGCAGAACGATGT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2508 ATTAAGCCGGCAGGCACCTATTGGAGGGCGACTTTCCAGCAGTTCAGAGAGCAGAACGATGT 2567

Qy 1561 ACATAGGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2568 ACATAGGGGCCTTCAAGAGGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGA 2627

Qy 1621 GACTGTACGAATATTTCTGACAGAGCAGCGCCTTTGGAAGGACTAGAGAAACCTCTACCAGGA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2628 GACTGTACGAATATTTCTGACAGAGCAGCGCCTTTGGAAGGACTAGAGAAACCTCTACCAGGA 2687

Qy 1681 GCCCAGAGAGCTGCCTCCTCCCAGAGAGAGAGCCCAGAATGTCACTCGGCTTCACGGAAAGCA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2688 GCCCAGAGAGCTGCCTCCTCCCAGAGAGAGAGCCCAGAATGTCACTCGGCTTCACGGAAAGCA 2747

Qy 1741 GGCTGAGGAGGTCAATACTGAGTGGGGGAAAAATTGAACCTGCACTCCGCTTCCGCTACTGGCCAGAG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2748 GGCTGAGGAGGTCAATACTGAGTGGGGGAAAAATTGAACCTGCACTCCGCTTCCGCTACTGGCCAGAG 2807

Qy 1801 AAAAATAGATGAGACCCCTTGAAAGACTTCAGGAGACTTCAAGAGGCCACGGATGAGCTGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2808 AAAAATAGATGAGACCCCTTGAAAGACTTCAGGAGACTTCAAGAGGCCACGGATGAGCTGGA 2867

Qy 1861 CCTCAAGTCGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2868 CCTCAAGTCGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCT 2927

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    April 4, 2004, 16:52:43 ; Search time 4197.89 Seconds
                                     (without alignments)
                                     17079.796 Million cell updates/sec

Title:        US-09-845-416-9_COPY_600_3000
Perfect score: 2401
Sequence:     1 acatgcattcaacatcgcca.........tggatcctcttgagcctgtca 2401

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      27513289 seqs, 1493109276 residues

Total number of hits satisfying chosen parameters:   55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  EST:*
    1: em_estba:*
    2: em_esthum:*
    3: em_estin:*
    4: em_estmu:*
    5: em_estcv:*
    6: em_estpl:*
    7: em_estro:*
    8: em_htc:*
    9: em_estl:*
   10: gb_es2:*
   11: gb_htc:*
   12: gb_est3:*
   13: gb_est4:*
   14: gb_est5:*
   15: em_estfun:*
   16: em_estom:*
   17: em_gss_hum:*
   18: em_gss_inv:*
   19: em_gss_pln:*
   20: em_gss_vrt:*
   21: em_gss_fun:*
   22: em_gss_mam:*
   23: em_gss_mus:*
   24: em_gss_pro:*
   25: em_gss_rod:*
   26: em_gss_phg:*
   27: em_gss_vrl:*
   28: gb_gss1:*
   29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Match	Query Match Length	DB	ID	Description
1		1093.8	45.6	3870	11	BC036103	BC036103 Homo sapi
2		994.4	41.4	5691	29	AY399453	AY399453 Homo sapi
3		864.6	36.0	3056	11	AK044536	AK044536 Mus muscu
4		786.6	32.8	5697	29	AY399455	AY399455 Mus muscu
5		607.8	25.3	5676	29	AY399454	AY399454 Pan trogl
6		605.4	25.2	834	12	BI729851	BI729851 603349511
7	c	593.4	24.7	595	14	CB177816	CB177816 is21c01.x
8		495.2	20.6	750	12	BI730168	BI730168 603349711
9		486.8	20.3	579	9	AL121550	AL121550 DKFZp762L
10		485.8	20.2	728	14	CB228986	CB228986 AGENCOURT
11		459.6	19.1	1047	14	CB850319	CB850319 MRA-0070
12		424.2	17.7	663	12	BM488464	BM488464 pgm2n.pk0
13		392.6	16.4	3753	11	AK081426	AK081426 Mus muscu
14		389	16.2	644	13	BU313510	BU313510 603540290
15		367.4	15.3	2874	29	AY408546	AY408546 Homo sapi
16		354.6	14.8	2874	29	AY408548	AY408548 Mus muscu
17	c	324.6	13.5	402	14	CB547284	CB547284 AMGNNUC:S
18		292.4	12.2	9915	29	AY407022	AY407022 Homo sapi
19		286.4	11.9	696	12	BJ075057	BJ075057 BJ075057
20		284.8	11.9	3051	11	BC036095	BC036095 Homo sapi
21		272.6	11.4	2874	29	AY408547	AY408547 Pan trogl
22		271.6	11.3	2334	11	BC011062	BC011062 Mus muscu
23		271.6	11.3	9691	29	AY407024	AY407024 Mus muscu
24		270.2	11.3	1122	14	CF109978	CF109978 Shultzomi
25		269.6	11.2	9096	29	AY407023	AY407023 Pan trogl
26		265.8	11.1	784	12	BG212445	BG212445 RST32032
27		260	10.8	826	14	CD656896	CD656896 AGENCOURT
28		257.8	10.7	1541	11	AK034383	AK034383 Mus muscu
29		256.8	10.7	630	9	AL855376	AL855376 AL855376
30		254	10.6	772	12	BI250598	BI250598 602993659
31		250	10.4	502	9	AL602076	AL602076 DKFZp213B
32		246	10.2	835	12	BI553820	BI553820 603190772
33		233.8	9.7	732	13	BU107880	BU107880 603109863
34		230.2	9.6	763	14	CB518960	CB518960 UI-M-GH0-
35		229.2	9.5	778	14	CB524596	CB524596 UI-M-FY0-
36		226	9.4	250	10	BF963618	BF963618 QV2-NN004
37		216.6	9.0	681	13	BU301653	BU301653 603609005
38		206.4	8.6	479	13	BQ304046	BQ304046 QV2-BT063
39		199.2	8.3	697	13	BU456556	BU456556 603771141
40		187.4	7.8	812	14	CF748008	CF748008 UI-M-HE0-
41	c	182.4	7.6	645	10	AW467977	AW467977 he29g08.x
42		179.6	7.5	466	14	CD549993	CD549993 B0305B01-
43		178.2	7.4	600	12	BI988528	BI988528 4012-24 M
44		174.6	7.3	851	13	BU201022	BU201022 603952191
45	c	172	7.2	646	12	BI289102	BI289102 UI-R-DK0-

ALIGNMENTS

```
RESULT 1
BC036103
LOCUS       BC036103                3870 bp    mRNA    linear   HTC 19-NOV-2003
DEFINITION  Homo sapiens dystrophin (muscular dystrophy, Duchenne and Becker
            types), mRNA (cDNA clone IMAGE:5274415), with apparent retained
            intron.
ACCESSION   BC036103
VERSION     BC036103.1  GI:23271310
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3870)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
```

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE   Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE 2 (bases 1 to 3870)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (31-JUL-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
          cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
          Toshiyuki and Piero Carninci (RIKEN)
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Sequencing Group at the Stanford Human Genome
          Center, Stanford University School of Medicine, Stanford, CA 94305
          Web site: http://www-shgc.stanford.edu
          Contact: (Dickson, Mark) mcd@paxil.stanford.edu
          Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
          R. M.

          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 48 Row: f Column: 1
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Similarity but not
          identity to protein
          This clone has the following problem: retained intron.
FEATURES   Location/Qualifiers
     source   1..3870
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5274415"
               /tissue_type="Brain, hippocampus"
               /clone_lib="NIH_MGC_95"
               /lab_host="DH10B"
               /note="Vector: pBluescript"
ORIGIN

Query Match   45.6%;   Score 1093.8;   DB 11;   Length 3870;
Best Local Similarity   95.2%;   Pred. No. 3.1e-238;
Matches 1128;   Conservative   0;   Mismatches   57;   Indels   0;   Gaps   0;

Qy   1 ACATGGCATTCAACACATCGCCAGATATCAATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   743 ACATGGCATTCAACACATCGCCAGATATCAATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGA   802

Qy   61 TGTTGATACCACCTATCCAGATAAGAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   803 TGTTGATACCACCTATCCAGATAAGAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   862

Qy   121 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   863 AGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC   922

Qy   181 ACCTAAAGTGACTAAAGAAGAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   923 ACCTAAAGTGACTAAAGAAGAGAACATTTCAGTTACATCATCAAATGCACTATTCTCAACA   982

Qy   241 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTCCCTAAGCCTCGATTCAA   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   983 GATCACGGTCAGTCTAGCACAGGGATATGAGAGAACTTTCTCCCTAAGCCTCGATTCAA   1042

Qy   301 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGACCCTACACGGAGCCCATT   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCCTGACCCTACACGGAGCCCATT   1102

Qy   361 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   1162

Qy   421 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGTTCTTTC   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGTTCTTTC   1222

Qy   481 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   1282

Qy   * 541 CCAGTTTCATACTCATGAGGGGTACAGCCCATCAGGGCCGGGGTGG   600
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1283 CCAGTTTCATACTCATGAGGGGTACAGATGGATTTGACAGCCCATCAGGGCCGGGGTGG   1342

Qy   601 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1343 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA   1402

Qy   661 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGGCCTCAGGGTAGC   720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1403 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   1462

Qy   721 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1463 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT   1522

Qy   781 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAAGAAAGGGTCAATTCTCCACTGGAGGAGA   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1523 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAAGAAAGGGTCAATTCTCCACTGGAGGAGA   1582

Qy   841 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACACAACATAAGGTGCTTCA   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1583 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACACAACATAAGGTGCTTCA   1642

Qy   901 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCCACTGTGGTGGTAGT   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1643 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCCACTGTGGTGGTAGT   1702

Qy   961 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGAAGAACAACTTAAGGTATTGGG   1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1703 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG   1762

Qy   1021 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGGTTCTTCA   1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1763 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGGTTCTTTTACAAGACAT   1822

Qy   1081 TCATAGATTACTGCAACAGTTCCCCTGGAAAGTTCTTGCCTGGCTTACAGA   1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1823 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCCTTTCAGA   1882

Qy   1141 AGCTGAAACAACTGCCAATGTCCTACCAGGATGCTACCCGTAAGGA   1185
           | |||
Db   1883 AAAGAGAGATGCAGTGAACAGAATTCACAACAACTGGCTTTAAAGA   1927

RESULT 2
AY399453
LOCUS   AY399453   5691 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Homo sapiens HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY399453
VERSION   AY399453.1   GI:39755442
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
Qy   620  AGTAAGCTGATTGGACACGAAGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAG  679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   604  AGTAAGCTGATTGGACACGAAGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAG  663

Qy   680  ATGAATCCTAAATTCCAGAATCCCAGAATGGGATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGC  739
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   664  ATGAATCCTAAATTCCAGAATCCCAGAATGGGATGCCTCAGGGTAGCTAGCATGGAAAAACAAAGC  723

Qy   740  AATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAGAGTTGAATGACTGG  799
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   724  AATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAGAGTTGAATGACTGG  783

Qy   800  CTAACAAAAACAGAAGAAAGAACAAGGAAATGGAGAAGAGCCTCTTGGACCTGATCTT  859
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   784  CTAACAAAAACAGAAGAAAGAACAACAAGGAAAATGGAGAAGAGCCTCTTGGACCTGATCTT  843

Qy   860  GAAGACCTAAAACGCCAGTACACACACATAGGTGCTTCAAGAAGATCTAGAACAAGAA  919
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   844  GAAGACCTAAAACGCCAGTACACAACACATAGGTGCTTCAAGAAGATCTAGAACAAGAA  903

Qy   920  CAAGTCAGGTCAATTCTCCACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGAT  979
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   904  CAAGTCAGGTCAATTCTCCACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGAT  963

Qy   980  CACCCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGCAAACATC  1039
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   964  CACCCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGGAGATCGATGGCAAACATC  1023

Qy  1040  TGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAG  1099
          |||||||||||||||||||||||||||||||||||||||||     ||||
Db  1024  TGTAGATGGACAGAGACCGCTGGGTTCTTTTTTACAAGACATCCTTCCAAATGGCAACGT  1083

Qy  1100  TTCCCCTGGACCTGG  1115
          |||||
Db  1084  CTTACTGAAGAACAGG  1099
```

RESULT 3
AK044536
LOCUS   AK044536   3056 bp   mRNA   linear   HTC 20-SEP-2003
DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library,
            clone:A930019F21 product:dystrophin, muscular dystrophy, full
            insert sequence.
ACCESSION   AK044536
VERSION   AK044536.1 GI:26090404
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
```
REFERENCE
  AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1  (bases 1 to 5691)
            Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 5691)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment
FEATURES             Location/Qualifiers
     source          1..5691
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1..>5691
                     /locus_tag="HCM0229"
ORIGIN

Query Match       41.4%;  Score 994.4;  DB 29;  Length 5691;
Best Local Similarity   95.0%;  Pred. No. 1.5e-215;
Matches 1041;  Conservative    0;  Mismatches   46;  Indels    9;  Gaps    1;

```
Qy    20  AGATATCAATTAGGCATAGAGAAACTACTGATCCTGAAGATGTTGATACCACCTATCCA  79
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    13  AGAAACCAAAAATGATGTGTAGTGTTAATGTGCTTACAGATGTTGATACCACCTATCCA  72

Qy    80  GATAAGAAGTCCATCTTAATGTACATCACATCACTCTTCCAAGTTTGCCTCAACAAGTG  139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    73  GATAAGAAGTCCATCTTAATGTACATCAACACATCACTCTTCCAAGTTTGCCTCAACAAGTG  132

Qy   140  AGCATTGAAGCCATCCAGGAAGTGGAAAATGTTGCCAAGGCCACCTAAAGTGACTAAAGAA  199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133  AGCATTGAAGCCATCCAGGAAGTCGAAAATGTTGCCAAGGCCACCTAAAGTGACTAAAGAA  192

Qy   200  GAACATTTCAGTTACATCATCAAATGCACTATTCTCAACACGACTCAGTCTCAGTCTAGCA  259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193  GAACATTTCAGTTACATCATCAAATGCACTATTCTCAACAGATCACGGTCAGTCAGTCAGCA  252

Qy   260  CAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAAGAGCTCAAGAGCTATGCCTACCACCAG  319
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   253  CAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAAGAGCTCAAGAGCTATGCCTACACACAG  312

Qy   320  GCTGCTTATGTCACCACCTCTGACCCTACAACGGGAGCCCATTCCTTCACAGGCATTGGAA  379
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   313  GCTGCTTATGTCACCACCTCTGACCCTACAACGGGAGCCCATTCCTTCACAGC--------  364

Qy   380  GCTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGT  439
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   365  -TCCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGT  423

Qy   440  TATCAAACAGCTTTAGAGAAGTATTATCGTGCGCTTCTTCTGCTGAGGACACATTGCAA  499
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424  TATCAAACAGCTTTAGAGAAGTATTATCGTGCGGCTTCTTCTGCTGAGGACACATTGCAA  483

Qy   500  GCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGACCAGTTCATCATGAG  559
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484  GCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGACCAGTTCATCATGAG  543

Qy   560  GGGTACACATGATGGATTTGACAGCCCATCAGGGGCCGGGTTGGTAATATTCTACAATTGGGA  619
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544  GGGTACACATGATGGATTTGACAGCCCATCAGGGGCCGGGTTGGTAATATTCTACAATTGGGA  603
```

```
JOURNAL      sequencing pipeline with 384 multicapillary sequencer
             Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE      20530913
PUBMED       11076861
REFERENCE    4
AUTHORS      The RIKEN Genome Exploration Research Group Phase II Team and the
             FANTOM Consortium.
TITLE        Functional annotation of a full-length mouse cDNA collection
JOURNAL      Nature 409, 685-690 (2001)
REFERENCE    5
AUTHORS      The FANTOM Consortium and the RIKEN Genome Exploration Research
             Group Phase I & II Team.
TITLE        Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
JOURNAL      Nature 420, 563-573 (2002)
REFERENCE    6  (bases 1 to 3056)
AUTHORS      Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
             Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
             Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
             Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,Y., Kasukawa,T.,
             Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
             Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
             Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
             Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
             Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
             Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
             Muramatsu,M. and Hayashizaki,Y.
TITLE        Direct Submission
JOURNAL      Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
COMMENT      cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Retina RNA was provided by Dr. Stefano Gustincich (Department of
             Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
             02115, USA) whose assistance is gratefully acknowledged. Please
             visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source      1..3056
                 /organism="Mus musculus"
                 /mol_type="mRNA"
                 /strain="C57BL/6J"
                 /db_xref="FANTOM_DB:A930019F21"
                 /db_xref="MGI:2409854"
                 /db_xref="taxon:10090"
                 /clone="A930019F21"
                 /tissue_type="retina"
                 /clone_lib="RIKEN full-length enriched mouse cDNA library"
                 /dev_stage="adult"
     misc_feature   1..3056
                 /note="dystrophin, muscular dystrophy (MGD|MGI:94909,
                 GB|NM_007868, evidence: BLASTN, 100%, match=999)"
ORIGIN
Query Match      36.0%;   Score 864.6; DB 11;  Length 3056;
Best Local Similarity  84.0%;   Pred. No. 5.1e-186;
Matches 1001;  Conservative  0;  Mismatches 184;  Indels  6;  Gaps  2;
```

```
Qy      1 ACATGCATTCAACATCGCCAGATATCAATTAGGCATAGAAACTACTCGATCCTGAAGA 60
            |||||||||||| ||||||||| |||||||| | |||||||||||||||||||||||||
Db    638 ACATGCCTTCAACATTGCAAAATGCCAGTTAGGCATAGAAAAACTTCTTGATCCTGAAGA 697

Qy     61 TGTTGATACCACCTATCCAGATAAGAAGTCCATCTTAATGTACATCACATCACTCTTTCA 120
            |||||    |||||||||||||  ||||||||||||  ||||| |||||| |||||| ||
Db    698 TGTTGCTACCACTTATCCAGACAAGAAGAGTCCATCTTAATGTACATCACATCACTCTTTCA 757

Qy    121 AGTTTTGCCTCAACAAGTGAGCATTGAGCCATCCAGGAAGTGGAAATGTTGCCAAGGCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    758 AGTTTTGCCACAACAAGTGAGCATTGAGCCATTCAAGAAGTGGAAATGTTGCCCAGGAC 817

Qy    181 ACCT---AAAGTGACTAAAGAACACATTTTCAGTTACATCATCAAATGCACTATTCTCA 237
            | ||    |||||| ||||| || ||| ||| |||||||||||||||||||||||||||
Db    818 ATCTTCAAAAGTAACTAGAGAAGAACAATTTTCAATTACATCACCAGATGCATTACTCTCA 877

Qy    238 ACAGATCACGGTCAGTCTAGCACGGGATATGAGAGACTTCT---TCCCCTAAGCCTCG 294
            |||||||||||||||| ||||||||||||||||||||||||      |||||||||||
Db    878 ACAGATCACAGTCAGTCTAGCACAGGGTCATGAACAAACTTCTTCATCTCCTAAGCCTCG 937

Qy    295 ATTCAAGAGCTATGCCTACACACAGCGTGCTTATGTCAAGGCTTATGTTCAACACGGAG 354
            ||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db    938 ATTCAAGAGTTATGCCTCCACACAGCGTGCTTATGTTGCCCACCTCTGACCCTACACAGAG 997

Qy    355 CCCATTTCCTTCACAGCATTGGAAGCTCCTGAGACAAGTGCATTGGCAGTTCATTGAT 414
            |||| ||||||||||||||||||||| |||||||||||||| |||||||||||||||||
Db    998 CCCCTATCCTTCACAGCATTTGGAAGCTCCCAGAGACAAGTGACACTTGCCAGTTCATTGAT 1057

Qy    415 GGAGAGTGAAGTAAACCTGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCT 474
            |||||||||||||||||||||||||||||||||| |||||||||| |||||||||||||
Db   1058 GGAGACGGAGTAAATCTGATAGTAGTAACCAAACTGCTTTAGAAGAAGTACTTTCATGGCT 1117

Qy    475 TCTTTCTGCTGAGGACACATTGCAAGGACAAGGAGAGATTTCAAATGATGTGGAAGAAGT 534
            |||||| |||||||| ||||| |||||||||||||| |||||||||||| ||| |||||
Db   1118 TCTTTCTGCCGAGGATACATTGCCGCACCAAGGAGAGATTTCAAATGATGTGTGAAGAAGT 1177

Qy    535 GAAAGACCAGTTTCATACTCCATGAGGGGTACATGAATGATTGCAGCCCCATCAGGGGCG 594
            |||||||||||||||||||||||||||||||||||| |||| |||||| ||||||||||
Db   1178 GAAGAACAGTTTCATGCTCATGAGGGATTCATGATGATCTGCATCTCATCAAGGACT 1237

Qy    595 GGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGA 654
            |||||||||| ||||| ||||||||| || |||||||| ||||||| |||||||||||
Db   1238 TGTTGGTAAGTGTTCACAGTTAGGAAGTCAACTAGTTGGAAAAGGGAAATTATCAGAAGA 1297

Qy    655 TGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCCTAAATTCAAGATGGAATGCCTCAG 714
            |||| || |||||||| ||||||| ||||| |||||||||||||||||||||| ||||||||||||
Db   1298 TGAAGAGTCGAAGTGCAAGAACAAATGAAT CTCCTAAATTCCTAAATTCAAGATGTCAG 1357

Qy    715 GGTAGCTAGCATGAAAAAACAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCA 774
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1358 GGTAGCTAGCATGAAAAAACAAGCAATTTACAACAAAGTTCTAATGGATCTCCAGAATCA 1417

Qy    775 GAAACTGAAGAGTTGAATGACTGGCTAACAACACAGAAGAAGAACCAAGAAAAATGA 834
            |||| |||||| ||| |||| |||| ||||||||||||||||||||||||||||||||
Db   1418 GAAATTAAAAGAACTAGATGACTGGTTAACAAAAACTGAAGAGAGAACTAAGAAAATGA 1477

Qy    835 GGAAGAGCCTCTTGGACCTGATCTTGAGACTCCAAGCCAAGTACAACACATAGGT 894
            |||||| ||| || |||||||| |||||||||||||||||||||||||||| ||||
Db   1478 GGAAGAGCCCTTTGGACCTGATCTTGAGACTCTAAAATGCCAAGTACAACAACATAGGT 1537

Qy    895 GCTTCAAGAGAGATCTAGAACAAGACAAGCAGGTCAGGTCAATTCTTCACTCACATGGTGGT 954
            ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   1538 GCTTCAAGAGAGATCTAGAACAAGACAAGCAGGTCAGGTCAACTCGTCTCACTCACATGGTAGT 1597

Qy    955 GGTAGTTGATGAATCTAGTGGAGATCACGGCAACTGCTGCTTTGGAAGAACAACTTAGGT 1014
            ||||||||||||| |||||||||||||||||| |||||||||||||||||| |||||||
Db   1598 AGTGGTTGATGAATCCAGCGGTGATCAACAGCGTGCTTTGGAAGAACAACTTAGGT 1657

Qy   1015 ATTGGGAGTCGATGGCAAACATCTGTAGATGACAGAAACCGCTGGGTTCTTTTTACA 1074
            ||||||||| || ||||||||||||||||||| |||| || |||| |||||||| ||||
Db   1658 ACTGGGAGATCGATGGGCAAATATCTGCAGATGACTGAAGATCGGGCTGGATTGTTTTTACA 1717

Qy   1075 AGACACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCT 1134
            |||| |||||||| | |||| |||||||||||||||||||||||||| | ||| ||||||
Db   1718 AGATATTCTTCTAAAATGGCAGCATTTTACTGAAGAACAGTGCCTTTTTTAGTACATGGCT 1777

Qy   1135 TACAGAAGCTGAAACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGA 1185
            | ||||||      | |||||
```

Db   1778 TTCAGAAAAAGAAGATGCAATGAAGAACATTCAGACAAGTGCTTTAAAGA 1828

RESULT 4
AY399455
LOCUS   AY399455   5697 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Mus musculus HCM0229 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY399455
VERSION   AY399455.1 GI:39755444
KEYWORDS   GSS.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 5697)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..5697
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene   <1..>5697
                     /locus_tag="HCM0229"
ORIGIN

Query Match   32.8%; Score 786.6; DB 29; Length 5697;
Best Local Similarity 84.2%; Pred. No. 3.5e-168;
Matches 927; Conservative   0; Mismatches 159; Indels 15; Gaps   3;

Qy   20 AGATATCAATTAGGCATAGAGAGAAACTACTGATCCTGAAGATGTTGATACCACTTATCCA 79
            | ||||||||||||||||||||||| |||||||||| |||||||||| |||||||||||
Db   13 AGAAACCAAAAAGTGATATGTGCTTAATGTGCTTACACTCTCTTCAAGTGCTACCACTTATCCA 72

Qy   80 GATAAGAAGTCCATCTTAATGTACATCATCATCACTCTTCGCCTCAACAAGTG 139
            || ||||||||||||||||||||||||||||| ||||||| ||||||||||
Db   73 GACAAGAAGTCCATCTTAATGTACATCATCATCACTCTTCAAGTTTTGCCCAACAAGTG 132

Qy   140 AGCATTGAAGCCATCCAGGAAGTGGAAATGTTGCCAGGCACCT---AAAGTGACTAAA 196
            ||||||||||||||||| ||||||||| |||||||||||| ||||||||||
Db   133 AGCATTGAAGCCATTCAAGAAGTGGAAATGTTGCCCCAGGACATCTTCAAAAGTAACTAGA 192

Qy   197 GAAGAACATTTTCAGTTACATCATCAAATGACTAATTCTCAACAGATCACCGGTCAGTCTA 256
            |||||||||||||||||||||||| |||| |||| |||||||||| |||||||||||
Db   193 GAAGAACATTTTCAATTACATCACCAGATCATTACTTCTTCAACAGATCACCAGTCAGTCTA 252

Qy   257 GCACAGGGATATGAGAGAACTTCT--TCCCCTAAGCCTCGATTCAAGAGCTATGCCTAC 313
            |||||||||||||||||||| | ||||||||||||||||||||| |||||||||||
Db   253 GCACAGGGCTATGAACAACAACTTCTTCATCTCCTAAGCCTCGATTCAAGAGTTATGCCTTC 312

Qy   314 ACACAGGCTGCTTATGTCACCACCCTCGACCCTCACACGAGCCCATTTCCTTCACAGCAT 373
            |||||||| |||||||||||||||| |||||||||||||||||||||||||||| |||
Db   313 ACACAGGCTGCTTATGTTGCCACCTGATTCCACAGCAGCCCCATCCTTCACAG--- 369

Qy   374 TTGGAAGCTCCTGAAGACAAGCAAGTCATTGGCAGTGCAGTGAGTAAGTGAAGTAAACCTG 433
            ||||| |||| || |||| ||||||| |||||||||||| ||||||||| |||||

Db   370 ------GCTCCCAGAGACAAGTCACTTGACAGTTCATTGATTGGACGGAGCGAGTAAATCTG 423

Qy   434 GACCGTTATCAAACAGCTTAGAGAGAGTATTATCGTGCTTCTTTCTGCTGAGGACACA 493
            || | || ||||||||| |||||||||||||||||||| ||||||| ||| |||| ||
Db   424 GATAGTTACCAAACTGCTTTAGAGAGAGTACTTCATCGGCTTCTTTCTGCCGAGGATACA 483

Qy   494 TTGCAAGCACAAGGAGATTTCTAATGATGGGAAGTGGTGAAAGACCAGTTTCATACT 553
            ||| | ||||| || |||||||||||||||| || ||||||| ||||||||||||||
Db   484 TTGCGAGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAAGACAGTTCATGCT 543

Qy   554 CATGAGGGTACATGATGATTTGACAGCCCATCTCATCAAGGACTTGTTGTAATATTCTACAA 613
            |||||||||||||||||||||| ||||| |||||||||| ||| |||| |||| ||||||
Db   544 CATGAGGGATTCATGATGATCTGACATCTGACATCTGTCTCATCAAGGACTTGTTGTAATGTTCTCACAG 603

Qy   614 TTGGGAAGTAAGCTGATTGGAACACAGAAAATTATCAGAAGTGAGAAACTGAAGTACAA 673
            ||||| ||||| ||| ||||| |||||| |||||||||||| || |||||| |||||||
Db   604 TTAGGAAGTCAACTAGTTGGAAATTATCAGAAAGGGAAATTATCAGAAGATGAAGAGCTGAAGTGCAA 663

Qy   674 GAGCAGATGAATCTCCTAAATTCAAGATGGGAAGTCCTCAGGTGAGCTAGCTAGCATGGAAAAA 733
            |||||| ||||||||| |||||||||| |||| || ||||| ||||| |||| |||||||||||
Db   664 GAACAAATGAATCTCCTAAATTCAAGATGGGAAGTGTCTCAGGGTAGCTAGCATGGAAAAA 723

Qy   734 CAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAACTGAAAGAGTTGAAT 793
            |||||||||||||||| || |||||||| |||||||||||||| |||| ||||||||
Db   724 CAAAGCAAATTACAACAAGTTCTAATGATCTCCAGAATCAGAAAATTAAAGAAGAACTAGAT 783

Qy   794 GACTGGCTAACAAAAACAGAGAAAGAACAACAGGAAAATGAGAGGAAGAGCCTCTTGGACCT 853
            |||||||||| ||||| |||||||| || |||| ||||||||||||||||||| ||||||
Db   784 GACTGGTTAACAAAAACTGAAGAGAGAACTAAGAAAATGAGAGAAGAGCCCCTTTGGACCT 843

Qy   854 GATCTTGAAGACCTAAAACGCCAAGTACAACACATAAGGTGCTTCAAGAGATCAGAA 913
            |||||||||||| |||| |||||||||||| |||||||||| ||||||||||||||||
Db   844 GATCTTGAAGATCTAAAATGCCAAGTACAACAACACATAAGGTGCTTCAAGAGATCTCAGAA 903

Qy   914 CAAGAACAAGTCAGGTCAATTCTCCACTCACATGGTGGTGTAGTTGATGAATCTAGT 973
            ||||||| |||| || ||||||| || || |||||| |||| ||||| |||||||
Db   904 CAGGAGCAGGTCAGGTCAACTGCGTCAACTCGCTCACATGGTAGTTGGTGATGAATCCAGC 963

Qy   974 GGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGAGATCGATGGGCA 1033
            |||||||| || ||||| |||| |||||||||||||||| |||| |||||||||||||
Db   964 GGTGATCATGCCAACAGCTGCTTTGGAAGAACAACTTAAGGTACTGGGAGATCGATGGGCA 1023

Qy 1034 AACATCTGTAGATGCAGCAGAGAGACGCGCTGGGTTCTTTTACAAGACACTCATAGATTACTG 1093
            |||| | |||| |||||||||| ||| || || ||| ||| ||||| ||||||| |
Db 1024 AATATCTGCAGATGCGACTGGACTGGAATGCGCTGGATTGTTTTACAAGAATATTCTTCTCTAAAATGG 1083

Qy 1094 CAACAGTTCCCCTGGACCTG 1114
            |||||||| || ||||| ||
Db 1084 CAGCATTTTACTGAAGAACAG 1104

```
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.M., Lu,F., Murphy,B.,
           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
           Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
           Rockville, MD 20850, USA
COMMENT    This sequence was made by sequencing genomic exons and ordering
           them based on alignment
FEATURES             Location/Qualifiers
     source          1..5676
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1..>5676
                     /locus_tag="HCM0229"
ORIGIN

Query Match        25.3%;     Score 607.8;  DB 29;  Length 5676;
Best Local Similarity  60.8%;     Pred. No. 2e-127;
Matches  646;  Conservative   0;  Mismatches  408;  Indels    9;  Gaps    1;

Qy     53 CCTGAAGATGTTGATACCACCTATCCAGATAAGAAGAAGTCCATCTTAATGTACATCACATCA 112
          |||||||||||||| ||||||||||||||||||| ||||||||||| |||||||||||||||
Db     31 CTTACAGATGTTGCTACCACCTATCCAGATAAGAAGAAGTCCATTTTAATGTACATCACATCA 90

Qy    113 CTCTTCCAAGTTTTGCTCAACAAGTGACATTGAAGCCATCCAGGAGTGGAAATGTTG 172
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     91 CTCTTCCAAGTTTTGCTCAACAAGTGAGCATTGAAGCCATCCAGGAGTGGAAATGTTG 150

Qy    173 CCAAGGCCACCTAAAGTGACTAAAGAAGAAGAACATTTTCAGTTACATCATCAAATGCACTAT 232
          ||||||||||||||||| ||||| |||||||||||| |||| |||||| |||||||||||||
Db    151 CCCAGGCCACCTAAAGTGACTAAAGAAGAAGAACATTTTCAATTACATCATCAAATGCACTAT 210

Qy    233 TCTCAACAGATCACCGGTCAGTCTAGCACAGGGGTATGAGAGAACTTCCTCCCCTAAGGCCT 292
          ||||||||
Db    211 TCTCAACAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 270

Qy    293 CGATTCAAGAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGG 352
          
Db    271 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 330

Qy    353 AGCCCATTCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTGGCAGTTCATTG 412
          
Db    331 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 390

Qy    413 ATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGG 472
          
Db    391 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN-------- 441

Qy    473 CTTCTTTCTGCTGAGGACACACATTGCAAGCACACAAGGAGGAGATTTCTAATGATGTGGAAGTG 532
          |||||| 
Db    442 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 501

Qy    533 GTGAAAGACCAGTTCATACTCATGAGAGGGTACATCATGATGGATTTGACAGCCATCAGGGC 592
          |||||||||
Db    502 NNNNNNNNNNNNNNNNNNNNNNNGGGTACATGATGGATNNNNNAGCCCATCAGGGC 561

Qy    593 CGGGTGGTATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAATTATCAGAA 652
          ||||||||||||     |||| |||||||||||||||||||||||||
Db    562 CGGGTGGTAATAANNNNNANTTGGAAGTNNNNTGGAAGTNNNNNTGGAACAGGAAAAATTATCAGAA 621

Qy    653 GATGAAGAAACTGAAGTACAAGAGCAGATGAATGAATCCTAAATTCAAGATGGGAATGCCTC 712
          |||||||||||||||||||||||||||||||||
Db    622 GATGAAGAAACTGAAGTACAAGAGCAGATGAATGAATCTNNNNNAATTCAAGATGGGAANNNNNN 681

Qy    713 AGGGTAGCTAGCATGGAAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCCCCAGAAT 772
          |||||||||||||||||          |||||||| ||||||||||||||
Db    682 NNNNNNNNNNNNNNNNNNNNNNNNNNTTACATAGAGTTTTAATGGATCCCCAGAAT 741

Qy    773 CAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAAACAGAAGGAAAATG 832
          |||||| ||||||||||||||||||||||  ||||||||||||||||||||||
Db    742 CAGAAACTGAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAAACAGAAGGAAAATG 801
```

```
Qy    833 GAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACAACATAAG 892
          |||||||||||||||||||||||||||||||||||     ||||||||||||||||||
Db    802 GAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAG 861

Qy    893 GTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTG 952
          |||||||        |||||||||||||||||||||||||||||||||||||||||||||
Db    862 NNNNNNCAAGAAGATTTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTG 921

Qy    953 GTGGTAGTTGATGAATCTAGTGGAGATCACGCACCTGCTCGTTGGAAGAAGAACAACTTAAG 1012
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 GTGGTAGTTGATGAANNNNANTGGAGACCACGCCACCTGCTCGTTGGAAGAAGAACAACTTAAG 981

Qy   1013 GTATTGGGAGATCGATGGCAAACATCTGTAGATGGACAGAGACACCGGTGGTTCTTTTA 1072
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 GTATTGGGAGATCGATGGGCAAACATCTGTAGGTGGACAGAAGAAGCCCGCTGGGTTCTTTTA 1041

Qy   1073 CAAGACACTCATAGATTACTGCAACAGTTCCCCTGGACCTGG 1115
          |||||||||
Db   1042 CAAGACATCCTTCTCAAATGGCACGTCTTTACTGAAGAACAGG 1084
```

```
RESULT 6
BI729851
LOCUS       BI729851                834 bp    mRNA    linear   EST 20-SEP-2001
DEFINITION  603349511F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
            mRNA sequence.
ACCESSION   BI729851
VERSION     BI729851.1  GI:15706864
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11908 row: e  column: 03
            High quality sequence stop: 796.
FEATURES             Location/Qualifiers
     source          1..834
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5357162"
                     /tissue_type="retina"
                     /clone_lib="NIH_MGC_94"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Query Match        25.2%;     Score 605.4;  DB 12;  Length 834;
Best Local Similarity  88.6%;     Pred. No. 4.4e-127;
Matches  690;  Conservative   0;  Mismatches  86;  Indels    3;  Gaps    3;

Qy   1028 TGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACACTCATAGA 1087
          |||||||    ||||||||||||||||||||||||||||||||||    |||||||||
Db     31 TGGGGAAACATTCATAAAAAGAGTAAGTAAGAGTAAGTGAGCAAGCAAGAGGCTGCTTTGGAAGAAACTCATAGA 90
```

```
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8857
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 448.

FEATURES             Location/Qualifiers
     source          1..595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6553129"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN

Query Match             24.7%;  Score 593.4;  DB 14;  Length 595;
Best Local Similarity   99.8%;  Pred. No. 2.2e-124;
Matches   594;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    215 CATCATCAAATGCACTATTCTCAACAGATCACGCTCAGTCTAGCACAGGGATATGAGAGA 274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    595 CATCATCAAATGCACTATTCTCAACAGATCACGCTCAGTCTAGCACAGGGATATGAGAGA 536

Qy    275 ACTTCTTCCCCTAAGCCTCGATTCAAGAGCTATGCCCACACACAGGCTGCTTATGTCACC 334
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    535 ACTTCTTCCCCTTAAGCCTCGATTCAAGAGCTATGCCCACACACAGGCTGCTTATGTCACC 476

Qy    335 ACCTCTGACCCTACACGGAGCCCATTTCCTTCACAGCATTGGAAGTCCTCAAGACACAAG 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    475 ACCTCTGACCCTACACGGAGCCCATTTCCTTCACAGCATTGGAAGTCCTCAAGACACAAG 416

Qy    395 TCATTGGCAGTTCATTGATGGAGGTGAAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTA 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    415 TCATTGGCAGTTCATTGATGGAGGTGAAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTA 356

Qy    455 GAAGAAGTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGACACAAGGAGATT 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    355 GAAGAAGTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGACACAAGGAGATT 296

Qy    515 TCTAATGATGTGGAAGTGGTGAAAGACCAGTTCATGCTCATGAGGGGTACATGATGATGAT 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    295 TCTAATGATGTGGAAGTGGTGAAAGACCAGTTCATGCTCATGAGGGGTACATGATGATGAT 236

Qy    575 TTGACAGCCCATCAGGGCCGGGCCGGGTTGGTAATATTCTACAATTGGAAGTAAGCTGATTGGA 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    235 TTGACAGCCCATCAGGGCCGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGATTGGA 176

Qy    635 ACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATGAATCCTAAAT 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    175 ACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATGAATCCTAAAT 116

Qy    695 TCAAGATGGGAATGCCTCAGGGTAGCTAGCATGAAAAACAAAGCAATTTACATAGAGTT 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    115 TCAAGATGGGAATGCCTCAGGGTAGCTAGCATGAAAAACAAAGCAATTTACATAGAGTT 56

Qy    755 TTAATGGATCCCAGAATCAGAACTGAAAGAGTTGAATGACTGGCTAACAAAA 809
```

```
RESULT 7
CB177816/c

LOCUS       CB177816             595 bp    mRNA    linear   EST 31-JAN-2003
DEFINITION  is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3'
            similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ;, mRNA sequence.
ACCESSION   CB177816
VERSION     CB177816.1  GI:28186206
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lenishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Bistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,I., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: is21c01.y1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium

Qy   1088 TTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTTCTTGCTGGCTTACGAAGAGCTGAA 1147
          |||||| |||||||||||| ||||||||||||||||||   ||  |||||||||| ||||
Db     91 TTACTGCGACAGTTCCCTGGACCTGGAAAAGTTTCTTTCTTCCTGGATTACGAAGCAGAA 150

Qy   1148 ACAACTGCCAATGTCCTACAGAGATGCTACACCGTAAGGAAAGGCTCCTAGAGACTCCAAG 1207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    151 ACAACTGCCAATGTCCTACAGACGCTTCCCGTAAGGAAGCAAGCTCCTAGAGAGACTCCC-AG 209

Qy   1208 GGAGTAAAGAGCTGATGAAACAATGGCAAGGCAGGTGAAATTGAAGCTCACACA 1267
          |||||||||||||||||||| || |||| ||||||||||||||||||||||||
Db    210 GGAGTCAGAGAGCTGATGAAACCATGGCAAGATCTCCAAGGAGAAATTGAAACTCACACA 269

Qy   1268 GATGTTTATCACAACCTGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCC 1327
          |||| ||| ||||||| ||||||||||||||||||||| |||||||||||||||||||
Db    270 GATATCTATCACAATCTTGATGAAAATGCCAAAAAATCCTGAGATCCCTGGAAGGTTCG 329

Qy   1328 GATGATGCAGTCAGTCCGTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAGTGAACTT 1387
          |||||||||||||||||  |||||||||||||| |||||||||||||||||||||||||||
Db    330 GATGAAGCACCCCTGTTACAAAGACGTTTGGATAAACATGAACTTCAAGTGGAGTGAACTT 389

Qy   1388 CGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT 1447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 CAGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGT 449

Qy   1448 CTGCACCTTTCTCTGCAGGAACTTCTGTGTGGCTGGCTACAGCTAGAAAGATGATGAATTAAGC 1507
          ||||||| |||||||||||||||||| ||||||||| ||||||||||||||||||||||||
Db    450 TTGCATCTTTCTCTTCAGGAACTTCTGTTGTGGCTGGCTACAGCTGAAAGATGATGAACTGAGC 509

Qy   1508 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAAGCAGATGACTAGG 1567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 CGTCAGGCACCCATCGGTGGTGGTGATTCCCAGCAGTTCAGAAGCAGAAGCAGATGATACTAGG 569

Qy   1568 GCCTTCAGAGGGAATTGAAAAACTAAAGAACCTGTAATCATCAGACTACTTCTTGAGACTGTA 1627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 GCCTTCAGAGGGAATTGAAAAACTAAAGAACCTGTAATCATCAGACTACTTCTGTCGGAGACTGTG 629

Qy   1628 CGAATATTTCTGACAGAGCAGAGCAGCCCTTGGACGAGGACTAGAGAAACTC-TACCAGGAGCCCAG 1686
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    630 AGAATATTTCTGACAGAGCAGAGCAGCTTTGGACGAGGACTAGAGAAACTCTTACCAGGAGCCCAG 689

Qy   1687 AGAGCTGCCTCCTGAGGAGAGAGCCCAGAATGTCACTCGGCTTCACTCGGCGAAAGCAGGCTGA 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 AGAACTGCCCCCTGAAGAAGAGAGCTCAGAATGTCAGTCGGCTTCCTCGGCTCCTACGGAAGCAGGCTGA 749

Qy   1747 GGAGGTCAATACTGAGTGGGGAAAAATTGAACCTGCACTC-CGCTGACTGGCAGAGAAAA 1804
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    750 AGAGGTCAACGCTGAATGGGACCAATTGAACCTGCGCTCAAGCTGATTGGCAGAGAAAA 808
```

```
Db   306 ATGGCCAAAAAATCCTGAGATCCCTGGAAGGTTCGGATGAAGCACCCCTGTTACAAAGAC 365
         |||||||||||| ||||||||||||||||||| |||||||| ||||  ||||||||||||
Qy  1353 GTTTGGATAACATGAACTTCAAGTGGAGTGGAGTGAACTTCGGAAAAAAGTCTCTCAACATTAGGT 1412
         |||||||||||||| || |||||||||||||||||||||||| ||||||||||||||||||||||
Db   366 GTTTGGATAACATGAATTTCAAGTGGAGTGGAGTGAACTTCAGAAAAAAGTCTCTCAACATTAGGT 425
Qy  1413 CCCATTGGAAGCCAGTTCTGACCAGTGGAGAGCGTCTGCACCTTTCTCTGCAGGAACTTC 1472
         ||||| ||||||| |||||||||||||||| ||||| || |||||||| ||||||||||
Db   426 CCCATTTGGAAGCAAGTTCTGACCAGTGGAAGCGTTGCATCTTTCTCTTCAGGAACTTC 485
Qy  1473 TGGTGTGGCTACAGTTGAAAGATGAATTAAGGCCGGCAGGCCATTGGAGGGCGACT 1532
         | ||| ||||||||||||||||||| |  | |||| ||||||| | | || ||| ||
Db   486 TTGTTTGGCTACAGTTGAAAGATGAGTGAGCGTCAGGCCACCCATGCGTGGTGATT 545
Qy  1533 TCCCAGCAGTTCAGAGAGCA-GAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACT 1591
         ||||||||||||||| | ||| ||  ||| | |||||||||||||||||||||||||||| |
Db   546 TCCCAGCAGTTCAGAGCACGAATGCAATGATATACATAGGGCCTTCAAGAGGGAATTGAAAAACT 605
Qy  1592 AAAGAACCTGTAATCATGAGTACTCTTGAGACTGGACTGTACGAATATTTCTGACAGAGCAGCCT 1651
         ||||||||||||||||||||||||||| | ||| | | || || |||||||||||||||||||||
Db   606 AAAGAACCTGTAATCATGAGTACTCTTGGGACTGTGAGAGATATTTCTGACAGAGCAGCCT 665
Qy  1652 TTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCC 1711
         |||||||||||||||||||||||||||||||||||||| |||| ||||||| ||||| |
Db   666 TTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGACTGCCCTCCTGAAGAAGAGCT 725
Qy  1712 CAGAATGTCACTCGGCTTCTACGA 1735
         |||||||||||||||||||  |||
Db   726 CAGAATGTCACTCGGCTTCCTACGA 749
```

RESULT 9
AL121550
LOCUS   AL121550   579 bp   mRNA   linear   EST 04-SEP-2003
DEFINITION DKFZp762L078_r1 762 (synonym: hmel2) Homo sapiens cDNA clone
            DKFZp762L078 5', mRNA sequence.
ACCESSION   AL121550
VERSION   AL121550.1 GI:5927551
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 579)
AUTHORS   Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
TITLE   EST (Ottenwaelder, et al.)
JOURNAL   Unpublished (1999)
COMMENT   Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by MediGenomix (Martinsried/Germany) within the cDNA
            sequencing consortium of the German Genome Project. No s1 sequence
            available.
            This clone (DKFZp762L078) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES   Location/Qualifiers
     source 1..579
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="DKFZp762L078"
             /tissue_type="melanoma (MeWo cell line)"
             /dev_stage="adult"
             /lab_host="DH10B"
             /clone_lib="762 (synonym: hmel2)"
             /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

```
Db    |||||||||||||||||||||||||||||||||||||||||||||||||||
   55 TTAATGGATCTTCAGAATCAGAGAACTGAAAGAGTTGAATGACTGCTAACAAAAA 1
```

RESULT 8
BI730168
LOCUS   BI730168   750 bp   mRNA   linear   EST 20-SEP-2001
DEFINITION 6034971F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357187 5',
            mRNA sequence.
ACCESSION   BI730168
VERSION   BI730168.1 GI:15707181
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 750)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11908 row: f column: 04
            High quality sequence stop: 747.
FEATURES   Location/Qualifiers
     source 1..750
             /organism="Mus musculus"
             /mol_type="mRNA"
             /db_xref="taxon:10090"
             /clone="IMAGE:5357187"
             /tissue_type="retina"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_94"
             /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
             Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
             Average insert size 3.3 kb. Library enriched for
             full-length clones and constructed by Life Technologies.
             Note: this is a NIH_MGC Library."

ORIGIN

Query Match   20.6%; Score 495.2; DB 12; Length 750;
Best Local Similarity 85.5%; Pred. No. 5.7e-102;
Matches 636; Conservative 0; Mismatches 93; Indels 15; Gaps 7;

```
Qy  1000 AGAACAACTTAAGGTATTGGGAGATCGATGGGCAAACATCGTAGATGGACAGCAGGACCG 1059
         ||||| | || || ||| |  |||   | ||| ||| ||  | ||| |||| || ||| |
Db    13 AGAATATCAATACTTCTTGGGAAACATTCACTAAACATGAGTAAGTGAGCAAGAGGCTG 72
Qy  1060 CTTGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAA 1119
         ||||||  |       |||| |||||||||||||||| || |||||| |||||||| | |
Db    73 CTTGGGAC------AGAAACTCATAGATTACTGCGACAGTTCCCTGGACCTGGAGGA 125
Qy  1120 GTTTCTTGCCTG-CTTACAGAAGCTGAAAC--AACTGCCAATGTCCTACAGGATGCTAC 1176
         |||||||  ||| ||| || ||||| ||||   | |||||||||||||||||| || ||
Db   126 GTTTCTTTCCTGGATTTACGGAAGCAGAAACAGAGCTGCCAATGTCCTACAGGACGCTTC 185
Qy  1177 CCG-TAAGGAAAGGCTCCTAGAAGACTCCAGGG--AGTAAAAGAGCTGATGAAACAATG 1233
         ||| ||||| ||||||||||| ||||||||||||   ||| | ||||| |||||| |||
Db   186 CCGCTAAGGAGAAGCTCCTAGCAAGACTCCAGGGGCAGTGAGAGACCTCAGAGAGGCTCATG 245
Qy  1234 GCAAGACCTCCAAGGTGAAA-TTGAAGCTCACACAGATGTTTATCACAACCTGATGAAA 1292
         ||||| ||||||||| ||||  ||||| ||||| ||||| | |||||||  || |||||
Db   246 GCAAGATCCTCCAAGGAGAAAGTTGAAACTCACCACAGATATCTATCACAATCTTGATGAAA 305
Qy  1293 ACAGCCAAAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCCTGTTACAAAGAC 1352
         ||||||||||  || |||||||||||||||||||  ||||| || || |||||||||||||
```

```
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_Rh_Ovl"
/note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."

ORIGIN

Query Match     20.3%;  Score 485.8;  DB 14;  Length 728;
Best Local Similarity  92.4%;  Pred. No. 7.8e-100;
Matches  511;  Conservative  0;  Mismatches  42;  Indels  0;  Gaps  0;

Qy    595 GGTTGGTATATTCTACAATTGGGAAGTAAGCTAGTTGGACAGGAAAATTATCAGAAGA 654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGTTGGTATATTCTACAATTGGGAAGTAAGCTAGTTGGACAGGAAAATTATCAGAAGA 60

Qy    655 TGAAGAAACTGAGTACAGAGCAGAGAATCCTAAATTCAAGATGGGAATGCCTCAG 714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGAAGAAACTGAGTACAGAGCAGAGAATCCTAAATTCAAGATGGGAATGCCTCAG 120

Qy    715 GGTAGCTAGCATGGAAGAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCA 774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGTAGCTAGCATGGAAGAAACAAAGCAATTACATAGAGTTTAATGGATCTCCAGAATCA 180

Qy    775 GAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGGAAGAAGAACAAGGAAAATGGA 834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAAACTGAAAGAGTTGAATGACTGGCTGACAAAAACAGAAGAAGAACAAGGAAAATGGA 240

Qy    835 GGAAGAGCCTCTTGGACCTGATCTTGAAGACATCAAGTCAAGTACAACAACATAAGGT 894
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAAAGAACCCTTGGACCTGATCTTGAAGACATCAAGTCAAGTACAACAACATAAGGT 300

Qy    895 GCTTCAAGAAGATCTAGAACAAGAACAAGAACAGTCAGGGTCAATTCTCTCACTCACATGGTGGT 954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCTTCAAGAGATCTAGAACAAGAACAAGAACAGTCAGGGTCAATTCTCTCACTCACATGGTGGT 360

Qy    955 GGTAGTTGATGAATCTAGTGGAGATCACGCACGCACTGCTCTTGGAAGAACAACTTAAGGT 1014
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGTAGTTGATGAATCTAGTGGAGATCACGCACGCACTGCTCTTGGAAGAACAACTTAAGGT 420

Qy   1015 ATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACA 1074
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACA 480

Qy   1075 AGACACTCATAGATTACTGCAACAGTTCCCCTGGACCTGGAAAAGTTCTTGCCTGGCT 1134
          ||||||||||||||||||||||||||||||||
Db    481 AGACATCCTTCAAATGCAACGTCTTACTGAAGAACAGTGCATTGCCTTTTTAGTGCATGGCT 540

Qy   1135 TACAGAAGTGAA 1147
          |||||||
Db    541 TTCAGAAAAGAA 553

RESULT 11
CB850319
LOCUS       CB850319                1047 bp    mRNA    linear    EST 01-SEP-2003
DEFINITION  MRA-0070 MOUSE ADULT RETINA Mus musculus cDNA 5', mRNA sequence.
ACCESSION   CB850319
VERSION     CB850319.1  GI:34380806
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1047)
AUTHORS     Yu,J., Farjo,R., MacNee,S.P., Baehr,W., Stambolian,D.E. and
            Swaroop,A.
TITLE       Annotation and analysis of 10,000 expressed sequence tags from
```

```
Query Match     20.3%;  Score 486.6;  DB 9;  Length 579;
Best Local Similarity  97.2%;  Pred. No. 4.8e-100;
Matches  495;  Conservative  0;  Mismatches  14;  Indels  0;  Gaps  0;

Qy   1236 AAGACCTCCAAGTGAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACA 1295
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db     71 AGGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGATGAAAACA 130

Qy   1296 GCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCGTTACAAAGACGTT 1355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131 GCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCGTTACAAAGACGTT 190

Qy   1356 TGGATAACATGAACTTCAAGTGGAGTGAACTTCTCAACAAAAGTCTCTCAACATTAGGTCCC 1415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 TGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACACATTAGGTCCC 250

Qy   1416 ATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGG 1475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 ATTTGGAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGG 310

Qy   1476 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTTC 1535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCACCTATTGGAGGCGACTTTC 370

Qy   1536 CAGCAGTTCAGAGAGCAGAAGCGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAG 1595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 CAGCAGTTCAGAGAGCAGAAGCGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAG 430

Qy   1596 AACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCCCTTTGG 1655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 AACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCCCTTTGG 490

Qy   1656 AAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTCGCCTCCTGAGGAGAGAGCCCAGA 1715
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 AAGGGCTAGAGAAACTCTACCAGGAGCCCAGAGAGTCGCCTCCTGAGGAGAGAGCCCCAG 550

Qy   1716 ATGTCACTCGGCTTCTACGAAAGCAGGCT 1744
          |||||||||||||||||||||||||||||
Db    551 AATGTCACTCGCTTCTACGAAAGCAGGCT 579

RESULT 10
CB228986
LOCUS       CB228986                 728 bp    mRNA    linear    EST 10-FEB-2003
DEFINITION  AGENCOURT 11492247 NICHD_Rh_Ovl Macaca mulatta cDNA clone
            IMAGE:6884820 5', mRNA sequence.
ACCESSION   CB228986
VERSION     CB228986.1  GI:28280564
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
ORGANISM    Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1 (bases 1 to 728)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapps-r@mail.nih.gov
            Tissue Procurement: Dr. Eliot Spindel
            cDNA Library Preparation: CLONTECH
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3135 row: e column: 11
            High quality sequence stop: 583.
FEATURES             Location/Qualifiers
     source          1..728
                     /organism="Macaca mulatta"
```

JOURNAL   developing mouse eye and adult retina
             Genome Biol. 4 (10), R65 (2003)
MEDLINE   22881944
PUBMED   14519200
COMMENT   Contact: Swaroop, A.
             Department of Ophthalmology and Visual Sciences
             Kellogg Eye Center, University of Michigan
             540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
             Tel: 734 615 2246
             Fax: 734 647 0228
             Email: swaroop@umich.edu.
FEATURES   Location/Qualifiers
    source   1..1047
             /organism="Mus musculus"
             /mol_type="mRNA"
             /db_xref="taxon:10090"
             /tissue_type="retina"
             /clone_lib="MOUSE ADULT RETINA"
             /note="Vector: pSPORT1"
ORIGIN

Query Match   19.1%;   Score 459.6; DB 14; Length 1047;
Best Local Similarity 80.3%; Pred. No. 8e-94;
Matches 578; Conservative   0; Mismatches 138; Indels   4; Gaps   4;

Qy   433 GGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGCCTTCTTTCTGCTGAGGACAC 492
            ||||| ||||| || | |||| |||| ||| |||||||| ||| ||| |||||||
Db   86 GGATAGTTACCAAACTGCTTTANAAGAAGTACTTTCATGCCTTCTTTCTGCCGAGGATAC 145

Qy   493 ATTGCAAGCACAAGGAGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTTCATAC 552
            |||| |||||| ||||| ||||||| ||| |||||||||||| |||| |||||| |||||
Db   146 ATTGCGGCACAAGGAGAGATTTCAAATGATGTTGAAGAAGTGAAAGAACAGTTTCATGC 205

Qy   553 TCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCCGGTTGGGTAATATTCTACA 612
            |||||||||| |||| ||||||||| ||||||||||||| ||||| || |||||||||
Db   206 TCATGAGGGATTCATGATGATCTGACATCTCATCAAGAGACTTGTTGGTAATGTTCTACA 265

Qy   613 ATTGGGAAGTAAGCTCGATTGGAACAGGAGAAAATTATCAGAAGATGAAGAAACTGAAGTACA 672
            |||||||||| |||| ||||||||| |||| ||||||||||||||||| ||||| |||||
Db   266 GTTAGGAAGTCAACTAGTTGGAAAAGGGAAATTATCAGAAGATGAAGAAGCTGAAGTGCA 325

Qy   673 AGAGCAGATGAATCCTAATTCCTAAATTCAAGATGGGAATGCCTCAGGGTAG-CTAGCAGTGAAAA 731
            || || ||||||| ||||||| ||||||||| |||||| ||||||| |||| ||||||||
Db   326 AGAACAAAATGAATCTCCTAAATTCAAGATGGGAATGTCTCAGGGTANCCTAGCATGAAAA 385

Qy   732 AACAAAGCAATTACATAGAGTTTTATGATCCTCCAGAATCAGAAACTGAAAGAGTTGA 791
            |||||| ||||| |||| |||||| ||||||| ||||||||| |||||||||||||
Db   386 AACAAAGCAAATTACACCAAAGTTCTAATGGATCCTCCAGAATCTCCAGAAATTAAAAGAACTAG 445

Qy   792 ATGACTGGCTAACAACAAGAAGAACAGAAGAAGAACAAGGAAAAATGGAGGAAGAGCCTCTTGGAC 851
            |||||||||||||||||||||| ||||||| ||||||||||||| | |||||||||||||||
Db   446 ATGACTGGTTTACAAAAACTGAAGAGAGAGAACTAANAAAAATGGAGGAAGAGCCCCTTGGAC 505

Qy   852 CTGATCTTGAAGACCTAAAAACGCCAAGTACAACAACACATAAGGTGC-TTCAAGAAGATCTA 910
            |||||||||| |||||||| || ||||||| || ||||| ||||||| |||||||||||
Db   506 CTGATCTTGAAGATCTAAAAATGCCAAGTACAACAACACATAAGGTGCTTCAAGAAGATCTA 565

Qy   911 GAACAAGAACAAGTCAGGGTCAATTCTCACTCCACATGGTGGTGAGTTGGATGAATCT 970
            |||||| ||||| ||| |||||||| |||||||| |||||| |||||||||| |||||
Db   566 GAACAGGAGAACAAGTCAGGGTCAACTCGCTCACATGNAGTAGTGGGTGATGAATCC 625

Qy   971 AGTGGAGATCACGCAACTGCTGCTTTGGAGAAACAAC-TTAAGGTATTGGGAGATCGATG 1029
            ||||| |||| ||| ||||| |||||||||||| | |||||| |||||||||| |||
Db   626 AGCGGGGATCATGCAACAGCTGCTTTGGAGANAACAACTTTAAGGNACTGGGAGATCGATG 685

Qy   1030 GGCAACAATCTG-TAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACACTCATAGAT 1088
            ||||| | | | ||| |||||||||| ||||||| || ||||||||||||||| |||
Db   686 GGCAAATATCTGCAAAAATGGACTGAAAAACCGTTGGATTGTTTTTACAAGAATTTTCTTCTA 745

Qy   1089 TACTGCAACAGTTCCCCTGGACCTTGGAAAAAGTTCTTGCCTGCTTACAGAGCTGAAA 1148
            |||||||| |||||||| ||| | |||||||||||| |||||||||||||||| |||||
Db   746 AAAATGGCAGCNTTTACTGAAAAANANNGCCTTTTATNCANGGNTTTTAAAAAAAAAAA 805

RESULT 12
BM488464
LOCUS   BM488464   663 bp   mRNA   linear   EST 07-FEB-2002
DEFINITION pgm2n.pk007.117 Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
            clone pgm2n.pk007.117 5' similar to sp|P11533|DMD_CHICK Dystrophin
            pir|S02041 dystrophin; muscle - chicken emb|CAA31746.1| (X13369)
            dystrophin (AA 1 - 3660) [Gallus gallus], mRNA sequence.
ACCESSION   BM488464
VERSION   BM488464.1 GI:18609395
KEYWORDS   EST.
SOURCE   Gallus gallus (chicken)
  ORGANISM Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 663)
  AUTHORS   Cogburn,L.A. and Monsonego-Ornan,E.
  TITLE   ESTs from Normalized Chicken Breast Muscle, Leg Muscle and
            Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
            Project
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES   Location/Qualifiers
    source 1..663
            /organism="Gallus gallus"
            /mol_type="mRNA"
            /strain="Commercial broiler and Ottawa Res. Centre
            Strains 90 & 21"
            /db_xref="taxon:9031"
            /clone="pgm2n.pk007.117"
            /sex="Male and Female"
            /tissue_type="Breast muscle, leg muscle and epiphyseal
            growth plate"
            /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
            ,11 weeks);growth plate(1d,7d,14d post-hatch)"
            /lab_host="E. coli EMDH10B"
            /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
            and Epiphyseal Growth Plate cDNA library (pgm2n)"
            /note="Vector: pCMVSPORT6; Library made from equivalent
            pools of total RNA isolated from each tissue (embryonic
            muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
            plate 33.3% of the final RNA pool). Single pass sequencing
            from 5'-end"
ORIGIN

Query Match   17.7%;   Score 424.2; DB 12; Length 663;
Best Local Similarity 78.1%; Pred. No. 8.3e-86;
Matches 518; Conservative   0; Mismatches 144; Indels   1; Gaps   1;

Qy   1219 GCTGATGAAACAATGCAAGACCTCCAAGTGAAATTGAAGCTCACACAGATGTTATCA 1278
            ||||| ||||||||||||| || ||||||||||||||||||||| ||||| || |||
Db   1 GCTCATGAAGCAGTGGCAGGATCTACAGGCAGAAATTGATGCACATACTGACATCT-TNN 59

Qy   1279 CAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 1338
            |||||||||||||||| || || ||||||||| |||||||||||| |||| ||||| ||
Db   60 CAACCTGGATGAAAACGGGCAGAAAATCCTGAGATCCCTGGAAGGCTCAGAGGATGCTGT 119

Qy   1339 CCTGGTTACAAAGAGACGTTTGGATAACAATGAAGCTTCAAGTGGAGTGGAACTTCGGAAAAAGTC 1398
            |||||||||||||||||||||||| |||||| ||||||| |||||| |||||||||||||| |||
Db   120 CCTGGTTGCAGAGACGTCTGATAACAATGAACCTTCAGATGGAGTGAGCTTAGGAGGAAATC 179

Qy   1399 TCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGGCTGCCACCTTTC 1458
            ||||| ||| ||||| |||||| |||| | ||||| |||||||| || ||| ||||| |
Db   180 TCTAAACATTAGATCTCATTTGGAGCCCAGAGCCACACAGACCAGTGGAAGCGTTTACATCTCTC 239

Qy   1459 TCTGCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCCAGGCACC 1518

```
Db        ||| |||| |||||| ||| |||||| ||||| | |||  | ||| ||||||
       240 TCTTCAGGAACTTTGGCATGCTGCTGCAATTGAAGGAGGAGGATGAATTAAAACAGCAAGCACC 299

2y    1519 TATTGGAGCGCGACTTTCAGCAGTTCAGAAGCAGAAGCAGCGATGTACATAGGGCCTTCAAGAG 1578
           |||| |||||||| |||| ||| |||  || |||||||||| ||| ||||||||||| |||||||
Db     300 CATTGGTGGAGATATTCCCACTGTGCAGAAGCAGAATGATGTTCATAGGACTTTCAAGAG 359

2y    1579 GGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCT 1638
           |||  |||||| || |||  ||| ||||| ||| ||||| |||||| ||||||||| ||
Db     360 GGAGCTGAAAACAAAAGAACCTGTTATCATGAATGAATGCTGCCGACTCTTCCT 419

2y    1639 GACAGAGCAGCCCTTTGGAAGGACTAGAGAAAACTCTACCAGGAGCCCAGAGAGCTGCCTCC 1698
           ||  |||| || |||||| | ||||| ||||| ||||||||| |||||||| ||||||||||
Db     420 GGCAGATCAACCAGTAGAGGGACTGGGAAAAGGTCTATCCAGAGAACCAAGAGACCTATCACC 479

2y    1699 TGAGGAGAGAGCCCAGAATGTCACTCGGCTTCACGAAAGACAGGCTGAGGAGGTCAATAC 1758
           ||||||||||| |||||||||||| ||| |||| |||||||| |||||||||||| ||||
Db     480 TGAGGAGAGGGCCCAGAATGTCACTAAAGTTCTCACTAAAGTTCTCCGAAGGCAAGCAGATGTGTCAGNAC 539

2y    1759 TGAGTGGGAAAAATTGAACCTGCCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCT 1818
           ||||| |||||  || |||||| |||  ||||| |||||||||||||| ||||||||||||
Db     540 TGAGTGGGATAAGCTAAATCTACCTTCGCTGATTGGCAAAAGAAGAAGATGATGCTCT 599

2y    1819 TGAAAGACTTCCAGGAACTTCAAGAGGCCACAGGACCTGGACCTGCCAAGC 1878
           ||||||||||||||| ||| |||  || | ||||  ||||| ||||||||
Db     600 TGAAAGACTGCAGGGTCTTCAGGAGGCAATGGGCAATGGACTNNNCCTGAAACTGCGCCAGGC 659

2y    1879 TGA 1881
           |||
Db     660 TGA 662
```

```
RESULT 13
AK081426
LOCUS       AK081426           3753 bp    mRNA    linear    HTC 20-SEP-2003
DEFINITION  Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
            library, clone:C130016K19 product:dystrophin related protein 2,
            full insert sequence.
ACCESSION   AK081426
VERSION     AK081426.1    GI:26349154
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   MEDLINE   99279253
   PUBMED    10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   MEDLINE   20499374
   PUBMED    11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   MEDLINE   20530913
   PUBMED    11076861
REFERENCE   4
```

```
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3753)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shirak,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..3753
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:C130016K19"
                     /db_xref="MGI:2413800"
                     /db_xref="taxon:10090"
                     /clone="C130016K19"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="16 days embryo"
     CDS             181..3054
                     /note="unnamed protein product; dystrophin related protein
                     2 (MGD|MGI:107432, GB|U43520, evidence: BLASTN, 99%,
                     match=807)
                     putative"
                     /codon_start=1
                     /protein_id="BAC38217.1"
                     /db_xref="GI:26349155"
                     /translation="MQPLVMQGCPYTLPRCHEWHAADRFHSSSLRNTCPQQVRAAV
                     TIPAPPWDGAGDPCLSPKLLNGTVGATGFLEPSAMNLCWNEIKKKSHNLRARLEAFSD
                     LSGKLQLPLREIIDWLSQKDEELSAQLPLQGVKETHAAFMEEVKSKGPYIS
                     SVLESAQAFLSOHPFEELEESQSESKDTSPRQRIQNLSRFWKQATVASELWEKLTAR
                     CVDQHRHIEHTLSHLLEIQGAMEELSSTLTQAEGVRATWEPIGDLFIDSLPEHIQAIK
                     LFKEFSPVKDGVKLVNDLAHQLAISDVHLSMENSRALEQINIRWKLQLVSVAERLKQ
                     LQDAHRDPGPGSQHFLSTSVQVPWERAFSPNKVPYYINHQAQTTCWDHPKMTELYQTL
                     ADLNNIKFSAYRTAMKLRRVQKALRLDLVTLITALEIFNEHDLQASEHVMVEVIHC
                     LTAIYERLEEERGILVNVPLCVDMSLWNLLNVFDSGRSGKMRALSFKTGIACLCGTEV
                     KEKLQYLFSQVANSGSOCDQRHLGALLHEAIQVPRQLGEVAAFGSGNVEPSVRSCFRF
                     STGKPVIRASQFLEWVNLEPQSMVLAVLHRVTIAEQVKGHQTKCSICROCPIKGFRYR
                     SLKQFNVDICQTCFLTGRASGKNLHYPIMBYYTPTTSSENMRDFATTLKNFRSKQY
                     FSKHPQRGYLPVQSVLESDCSETPASSPMLPHADTHSRIEHFASRLAEMESQNCSFFN
                     DSLSPDDSIDEDCYLLRHSSPITDREPAFGQQAPCSMATESKGELEXILAHLEDENRI"
```

```
            LQGELRRLRKWQHERAAEAPTLVEGSAEATPDHRNEBLLAEARILRQHKSRLETRMQIL
            EDHNKQLESQLQRLRELLLQPPESDGNGSAGSSLASSPRQSEGSHPREKGQTTPDTE
            VADDVGSKSQDVSLCLEDIMEKLRHAFPSVRSSDVTANTLLAS"
polyA_signal    3731..3736
                /note="putative"
polyA_site      3753
                /note="putative"
ORIGIN

Query Match     16.4%;    Score 392.6;  DB 11;  Length 3753;
Best Local Similarity   61.3%;   Pred. No. 2e-78;
Matches 632;   Conservative  0;   Mismatches 399;   Indels  0;   Gaps  0;

Qy  1363  CATGAACTTCAGTGGAGTGAACTTCGGAAAAAGTCTCAACATTAGGTCCCATTTGA 1422
Db   414  CATGAATCTGTGTGGAATGAAATAAAAAGAGAGTCTCACAACCTCCGGGGGCTCGCCTAGA 473

Qy  1423  AGCCAGTTCTGACCAGTGGAAGCGTTCGCACCTTTCTCTGCACCTTCTCTGGTGTGGCT 1482
Db   474  GGCTTCTTCAGACCTCAGTGGAAAACTTCAGCTCCCGCTCCGAGAGATTATTGACTGGCT 533

Qy  1483  ACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGT 1542
Db   534  CAGCCAAAAGATGAGGAGTTGTCAGCTCAGCTCCCTTGCAAGGGGATGTGGCCCTGGT 593

Qy  1543  TCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGGGGAATTGAAAACTAAAGAACCTGT 1602
Db   594  ACAACAGGAGAGGAGACACATGCAGCCTTCATGGGAGAGTCAGTCAGCAGCCGTTTGGAAGGACT 653

Qy  1603  AATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACT 1662
Db   654  CATCTCCTCTGTGCTTGAATCCGCCCAGGGTTTCTGTCGTCCCCAGCATCCATTTGAAGAATT 713

Qy  1663  AGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGAGCCCAGAAGTCAC 1722
Db   714  AGAGGAGTCCAATCTGAGAGTAAAGATACACCTCCCCCAGACAGCGGATTCAAAACCTTAG 773

Qy  1723  TCGGCTTCTACGAAAGCAGGGTGAGGAGGTCAATACTGAGTGGAAAAATTGAACCTGCA 1782
Db   774  TCGCTTTGTATGGAAGCAGCAGGCAACAGTGGCCAGTGACGAGTCGTGTGGGAGAAGCTGACGCCCG 833

Qy  1783  CTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTGAAAGACTCCAGGAACTTCAAGA 1842
Db   834  CTGTGTAGATCAGCATCGCCACATTGAGACTCTCTGGAACATCTATTGGAGATCCAAGG 893

Qy  1843  GGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGTGATCAAGGGATCCTGGCA 1902
Db   894  GGCAATGGAGGAATTGAGCAGTACTTGACCCAAGCAGCAGGGAGTCCGAGCCACGTGGGA 953

Qy  1903  GCCCGTGGGCGATCTCCATTGACTCTCTCCATTGACTCCTCCAAGATCACCTCGAAAGTCAAGGCACT 1962
Db   954  GCCCATAGGAGAGATCTCTTTATTGATTCCCTCCCAGAGCATATCCAAGCCATCAAGCATCAAGCTATT 1013

Qy  1963  TCGAGGAGAAATTGCGCCTCTGAAAGAGAAACGTGAGCCACGCACGTCAATGACCTTGCTCGCCA 2022
Db  1014  CAAAGAGAATTCTCTCCGTGTGAAAGATGGGGTGAAGATGGGGTGAGTGAAGTAGTGAATGATCTGGCCCACCA 1073

Qy  2023  GCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCCAGCACTCTGAAGACACTCGAA 2082
Db  1074  GCTTGCCATCTCTGATGTGCACTTGTCATGTCAATGGAGAATTCCAGGGGCTCTGGAACAGATCAA 1133

Qy  2083  CACCAGATGGAAGCTTCGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGC 2142
Db  1134  CATCCGGTGGAAACAGCTGCCAGTGTCAGTGGCTGCAGAGGCTTAAGCAACTCCAGGATGC 1193

Qy  2143  CCACAGGGACTTTGGTCAGCATCTCAGCACTTCTTCCACGTCTGTCCAGGGTCCCTG 2202
Db  1194  CCACCGGGACTTTTGGGGCCTGGACACAGGCCCAAACAAAATGTCCCATTTCCCTG 1253

Qy  2203  GGGAGAGAGCCACCATCTGCCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAAC 2262
Db  1254  GGAAAGAGCATTTTCTCCCAATAAAGTCCCCTACTACATCAACCACCACCAGGCTCAGACCAC 1313
```

```
Qy  2263  TTGCTGGGACCATCCCAAAATGACAGAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGT 2322
Db  1314  ATGCTGGGACCATCCTAAGATGACTGAGTTATACCAAACCCTAGCCGATCTGAACAACAT 1373

Qy  2323  CAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGAACTCGAGAGGCCCCTTGCTT 2382
Db  1374  TAAGTTCTCAGCTTACCGCACTGCCACTGGCAGAGCTCCGCAGAGTCCAGAAGGCCCTGCGGCCT 1433

Qy  2383  GGATCTCTTGA 2393
Db  1434  GGATCTCGTAA 1444
```

```
RESULT 14
BU313510
LOCUS       BU313510          644 bp    mRNA       linear   EST 28-NOV-2002
DEFINITION  603540290F1 CSEQCHN61 Gallus gallus cDNA clone ChEST508c24 5', mRNA
            sequence.
ACCESSION   BU313510
VERSION     BU313510.1      GI:25821511
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 644)
AUTHORS     Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE       A Comprehensive Collection of Chicken cDNAs
JOURNAL     Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE     22335534
PUBMED      12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1..644
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Compton Line 15I"
                     /db_xref="taxon:9031"
                     /clone="ChEST508c24"
                     /sex="Female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN61"
                     /note="Organ: heart; Vector: pBluescript II KS(+); Site_1:
                     EcoRI; Site_2: NotI; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
ORIGIN

Query Match     16.2%;    Score 389;  DB 13;  Length 644;
Best Local Similarity   75.5%;   Pred. No. 8.7e-78;
Matches 482;   Conservative  0;   Mismatches 156;   Indels  0;   Gaps  0;

Qy  1425  CCAGTTGACCAGTGAGCCGTCGCACCTTTCTCTGCACCTTCTCGTGTGGTGGCTAC 1484
```

```
Db      7 CGAGCACAGACCAGTGGAAGCGTTTACATCTCTCTTCAGGAACTTTTGGCATGGCTGC   66
              |||||||||||||||| |||||||| | ||| ||||||||| |||| |||

Qy   1485 AGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTC 1544
              |||||||||||||| |||| | ||||||||| ||||| ||||| |||||
Db     67 AATTGAAGGAGGATGAATTAAAACAGCAAGCACCCATTGGTGGAGATATTCCCACTGTGC  126

Qy   1545 AGAAGCGAACGATGATGTTCATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAA 1604
              |||||||||||| || |||||||| |||||||| ||||||||||| ||||||||||
Db    127 AGAAGCGAACGAATGATGTTCATAGGACGCTGAAAAACAAAGAACCTGTTA  186

Qy   1605 TCATGAGTACTCTTGAGACTGTACGAATATTTCGACAGAGCAGCCTTTGGAAGGACTAG 1664
              |||||||||||||||||||| |||| ||||| ||| |||||||||||||||||||||
Db    187 TCATGAATGCACTTGAGACTGTGCGACTCTTCCTGGCAGATCAACCAGTAGAGGGACTGG  246

Qy   1665 AGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGCCCAGAATGTCACTC 1724
              |||| || |||||||||||| ||||||||||||||||||||||||| ||||||||
Db    247 AAAAGTCTATCCAGGAGACCACCAGAGACCTATCACCTGAGGAGGAGGGCCCAGAATGTCACTA  306

Qy   1725 GGCTTCTACGAAAGCAGCAGGGTGAGGAGTCAATACTGAGTGGGAAAAAATTGAACCTGCACT 1784
              ||||||||||||||||||||| |||||||||| ||||| ||| ||||||| |||||||||
Db    307 AAGTTCTCCGAAGGCAAGCAGCAGATGATGTCAGAACTGAGTGGGATAAGCTAAATCTACGTT  366

Qy   1785 CCGCTGACTGCAGAGAAAAATAGATGAGGACCCTTGAAAGACTCCAGGAACTTCAAGAGG 1844
              ||||| || ||||||||||||||| |||| ||||||| || |||||||||||||||||
Db    367 CTGCTGATTGGCAAAAAGAAGATAGATGATGCTCTTGAAGAGACTCGCAGGTCTTCAGGAGG  426

Qy   1845 CCACGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGTGATCAAGGGATCCTGGGCAGC 1904
              ||||||||||| |||||||||||| ||| ||||||| |||||||||||||||||||||
Db    427 CAATGATGAACTAGACCTGGGAAAACTGGCCGCCAGGCTGAAGCATTCAAGGGATCCTGGGCAGC  486

Qy   1905 CCGTGGGGATCTCCATTGACTCTCCAAGATCACCTGAGAAAGTCAAGGCACTTC 1964
              |||||| |||| ||||||||||||||||||||||| ||||||||||||||||||
Db    487 CAGTGGGGATCGTCTGATAGACTCTCTGCAGGATCACCTTAGAAAAAAGTCAAGGTTTATC  546

Qy   1965 GAGAGAAATTGCGCCTTGAAAAGAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGC 2024
              ||| |||||||| ||| |||||||||| ||| |||||| ||||||||||| || ||||
Db    547 GAGCAGAAATGGTGCCCTTANAGAGAGGTGCATCAAGTCAATGAGCTGGCTCACCGGT  606

Qy   2025 TTACCACTTTGGGATTCAGCTCTCACCGTATAACCTC 2062
              |||||||| ||||||||||||| |||||||||||||
Db    607 TCGCTCCCCTGATATTCAGTTCTCCCCATACACTCTC  644
```

COMMENT   This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES   Location/Qualifiers

source   1..2874
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"

gene   <1..>2874
     /gene="DRP2"
     /locus_tag="HCM3260"

ORIGIN

Query Match   15.3%;   Score 367.4;   DB 29;   Length 2874;
Best Local Similarity   57.5%;
Matches 593;   Conservative   0;   Mismatches 438;   Indels   0;   Gaps   0;

Pred. No. 1e-72;

```
Qy   1363 CATGAACTTCAAGTGGAGTGAACTTCAACATTAGTCCATTGGA 1422
              ||||||||||||||||||||||||| ||||| ||||||||||
Db    234 CATGAATCTGTTGGAATGAAATAAAAAGAGAGTCTCACAACCTCCGCGCTCGCCTAGA  293

Qy   1423 AGCCAGTTCTGACCAGTGGAGAGCGTCTGCACCTTCTCGTGTGTGCT 1482
              |||| || ||||||||||| ||||||| |||| ||| |||||||||
Db    294 GGCCTTCCAGACCACCAGTGGAAAGCTTCAGCTCCCTCTTCAAGAGATTATTGACTGGCT  353

Qy   1483 ACAGTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCGACTTTCCAGCAGT 1542
              |||||||||||||||||| ||||||||||||||||||| |||| ||||| ||||||||
Db    354 CAGCCAAAAGGATGAGGAGTTGTCAGCTGCGCCCCTACAGGGGGATGGTGGCCCTCGT  413

Qy   1543 TCAGAAGCAGAACGATGTACATAGGGCCTCAAGAGGGAATTGAAAACTAAAGAAACCTGT 1602
              |||||||||||||||| |||||||||||||||||||||||| ||||||| |||||||||
Db    414 GCAACAGGAGAAGGAGACACATGCGGGCCTTTATGGAAGAAGTCAAGTCTCGGGCCCCTA  473

Qy   1603 AATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCCTTTGGAAGGACT 1662
              |||||||| || |||||||||| ||||| |||||| || |||||| ||||||||||
Db    474 CATCTATTCTGTGCGAGTCAGCTCAGGCCTTCAGGGACCCCATTTGAGGAGTT  533

Qy   1663 AGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAGCCCAGAATGTCAC 1722
              |||||||||||||||||||||||||||| ||||| |||||||||||||||| |||||||
Db    534 AGAGGAGCCTCATTCTGAGAGCAAAGATACCTCCCGAAACAGCGGATCCAGAATCTCAG  593

Qy   1723 TCGGCTTCTACGAAAGCAGCAGGTGAGGAGGTCAATACTGAGTGGGAAAAAATTGAACCTGCA 1782
              ||||| |||||| ||||||||| ||||||| |||||||||||||||||||||| ||||||
Db    594 CCGCTTTGTATGAAGCAGCAGCAGCCGACAGGTGGAACGTGGCCAGTGAACTGTGTGGGAGAAGTTGACAGCCCG  653

Qy   1783 CTCCCGCTGACTGGCAGAGAAAAAATAGATGAGATGGCGGACTCAAGAACTCCAAGA 1842
              ||||| |||| ||||||||||||| ||| ||| || |||||| ||| ||| |||||
Db    654 CTGTGTGGACCAGCACCGTCACATTGAGCGGACTCGGAGCAGCTGTTGGAGATTCAAGG  713

Qy   1843 GGCCACGGATGAGCTGGACCTGCAAGGTGCGCCCAAGCTGAGGTGATCAAGGATCCTGCA 1902
              ||| |||| |||||||| | ||||||||||| |||||| ||||| |||| ||||||||||
Db    714 GGCAATGGAGGAACTAAGCACTACTCTGAGCCAAGCTGAGGGAGTCCGAGCCACTTGGA  773

Qy   1903 GCCCGTGGCGGCGGATCTCCTCATTGACTCTCTCCAAGATCACCTCCGAGAAAGTCAAGGCACT 1962
              |||||||||| ||||||||| |||||| ||||||||||||||||||| ||| ||||||||||||
Db    774 GCCCATTGGGGATCTTCTTCATTGATTCATTGATTCACTGAGGACACATCCAGGTATTAAGCTGTT  833

Qy   1963 TCGAGGAGAAATTGCGCCTCTGAAAGAAGAACGTGAGCCACGTCAATGACCTTGCTCGCCA 2022
              ||| |||||| |||||| | ||||||||||| |||| ||| ||| ||||| |||||||| |
Db    834 CAAAGAAGAATTCTCCCCCATGAAAGAAGATGGAGTAAAGTTGGTGAATGATCTGGCCCACCA  893

Qy   2023 GCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTTCTTCAGCACTTCTCAAGGACCTGAA 2082
              ||| ||||||||||| |||||||||||||||||| || |||||||| |||||||||||||||||
Db    894 ACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAATTCCCAGGCCCTGGAACAGACATCAA  953

Qy   2083 CACCAGATGGAAGTTCTGGCCGTGGCCGTGGCGGCCGTCAGGCGAGCTGCATGAAGC 2142
              |||||||||||||| ||||||| |||||||| |||| ||||||||||||||||||||
Db    954 CGTCCGATGGAAACACTACAGGCGTCAGGTCAGTAGTGAGAGGCTTCAGGATGC 1013

Qy   2143 CCACAGGGGACTTTGGTCCAGCATCAGCACTTTCTTCCACGTGTCCAGGTTCCCTG 2202
              |||||| |||||||||||||||| |||| ||||||| ||| |||||||||||||||
Db   1014 CCACCGGGGACTTTGGGCTTGGGCCACACAACTTTCTCTCCTCTCTGTCTCCAGTTCCCTG 1073

Qy   2203 GGAGAGAGCGGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAC 2262
              ||||||||||||||||||||| |||| |||||||||||||||||||||||||||||||||||
```

RESULT 15

AY408546

LOCUS   AY408546   2874 bp   DNA   linear   GSS 15-DEC-2003
DEFINITION   Homo sapiens DRP2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION   AY408546
VERSION   AY408546.1   GI:39764517
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2874)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
     Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
     Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
     Adams,M.D. and Cargill,M.
TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE   2 (bases 1 to 2874)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
     Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
     Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
     Adams,M.D. and Cargill,M.
TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

```
                 |||| ||||| ||| || || ||||| ||||| |||||||| |||||
Db        1074 GGAAAGAGCAATTTCACCCAATAAAGTTCCCTACTACATCAANNNNNNNNNNNNNNN 1133

Qy        2263 TTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGT 2322
                                                                 ||||| |||| |||  |
Db        1134 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCTGATCTGAACAACAT 1193

Qy        2323 CAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAGGCCCTTTGCTT 2382
                    | ||||||||||| |       |||||||||||| ||| |||| ||||  ||||
Db        1194 TAAGTTCTCAGCTTATCGCACTGCCATGAAACTCCGCAGAGTCCAGAAAGCCCTGCGCTT 1253

Qy        2383 GGATCTCTTGA 2393
               ||| || ||-|
Db        1254 GGACCTGGTAA 1264
```

Search completed: April 5, 2004, 08:15:07
Job time : 4201.22 secs

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on:      April  4, 2004, 16:58:13 ; Search time 63.0939 Seconds
                                        (without alignments)
                                        11443.127 Million cell updates/sec

Title:          US-09-845-416-10_COPY_1800_3100
Perfect score:  1301
Sequence:       1 cgactttccagcagttcaga.........ccggagctgcttccaattg 1301

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       682709 seqs, 277745446 residues

Total number of hits satisfying chosen parameters:         1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued_Patents_NA:*
            1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
          score greater than or equal to the score of the result being printed,
          and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1301	100.0	5952	4	US-09-687-875A-1	Sequence 1, Appli
2	1290	99.2	13977	4	US-09-484-970B-60	Sequence 60, Appl
c 3	1135.2	87.3	13307	3	US-09-484-970B-60	Sequence 60, Appl
c 4	1135.2	87.3	19307	3	US-09-427-048A-10	Sequence 10, Appl
5	553.8	42.6	6045	4	US-09-091-501B-7	Sequence 7, Appli
6	553.8	42.6	10320	4	US-09-091-501B-9	Sequence 9, Appli
7	540.2	41.5	3915	4	US-09-976-594-93	Sequence 93, Appl
c 8	.50	3.8	7218	1	US-08-232-463-14	Sequence 14, Appl
9	44.2	3.4	2574	4	US-09-668-313A-10	Sequence 10, Appl
10	42.8	3.3	1690	4	US-09-620-312D-69	Sequence 69, Appl
11	42.8	3.3	7812	1	US-09-368-590-1	Sequence 1, Appli
12	38.6	3.0	1995	2	US-08-425-069-3	Sequence 3, Appli
13	38.6	3.0	1995	2	US-08-317-844B-3	Sequence 3, Appli
14	38.4	3.0	7672	4	US-09-220-132-24	Sequence 24, Appl
15	38.2	2.9	428	4	US-09-668-313A-3	Sequence 3, Appli
16	38.2	2.9	4439	4	US-09-668-313A-17	Sequence 17, Appl
17	34.6	2.7	474	4	US-09-621-976-18033	Sequence 18033, A
c 18	34.6	2.7	2277	1	US-08-676-967-5	Sequence 5, Appli
c 19	34.6	2.7	2277	1	US-08-676-974-5	Sequence 5, Appli
c 20	34.6	2.7	2277	2	US-09-098-487-5	Sequence 5, Appli
21	34.4	2.6	1047	4	US-09-671-950-1	Sequence 1, Appli
22	34.4	2.6	1047	4	US-09-671-950-3	Sequence 3, Appli
23	34.4	2.6	1047	4	US-09-671-950-5	Sequence 5, Appli
24	34.4	2.6	1047	4	US-09-671-950-7	Sequence 7, Appli
25	34.4	2.6	1047	4	US-09-671-950-9	Sequence 9, Appli
26	34.4	2.6	1047	4	US-09-671-950-11	Sequence 11, Appl
27	34.4	2.6	1047	4	US-09-671-950-13	Sequence 13, Appl
28	34.2	2.6	750	4	US-08-961-527-370	Sequence 370, App
29	34.2	2.6	1620	4	US-08-858-207A-56	Sequence 56, Appl
30	34.2	2.6	2800	2	US-08-874-138-1	Sequence 1, Appli
31	34.2	2.6	2800	2	US-08-874-138-5	Sequence 5, Appli
32	34.2	2.6	2800	3	US-08-879-941-1	Sequence 1, Appli
33	34.2	2.6	2800	4	US-08-879-941-3	Sequence 3, Appli
34	34.2	2.6	2800	4	US-09-747-116-1	Sequence 1, Appli
35	34.2	2.6	32768	4	US-09-747-116-3	Sequence 3, Appli
36	34.2	2.6	32768	4	US-08-961-527-71	Sequence 71, Appl
37	34	2.6	648	4	US-09-252-991A-10033	Sequence 10033, A
c 38	34	2.6	762	4	US-09-252-991A-9821	Sequence 9821, Ap
39	34	2.6	1644	4	US-09-252-991A-10161	Sequence 10161, A
c 40	34	2.6	2235	3	US-09-153-804-2	Sequence 2, Appli
c 41	33.8	2.6	505	4	US-09-621-976-15639	Sequence 15639, A
42	33.8	2.6	1603	1	US-08-625-209A-1	Sequence 1, Appli
43	33.8	2.6	3489	2	US-08-728-323A-1	Sequence 1, Appli
44	33.8	2.6	3489	4	US-09-299-568-1	Sequence 1, Appli
45	33.8	2.6	3489	4	US-09-410-399-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
;  APPLICANT: Xiao, Xiao
;  APPLICANT: Liu, Paul
;  TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
;  FILE REFERENCE: 00792
;  CURRENT APPLICATION NUMBER: US/09/687,875A
;  CURRENT FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: 60/158,868
;  PRIOR FILING DATE: 1999-10-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;  LENGTH: 5952
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (2897)..(2898)
;  OTHER INFORMATION: S4 junction site
;  NAME/KEY: misc_feature
;  LOCATION: (3198)..(3199)
;  OTHER INFORMATION: S2 junction site
US-09-687-875A-1

  Query Match         100.0%;  Score 1301;  DB 4;  Length 5952;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1301;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy           1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3402 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3461

Qy          61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGGAATATTTCTGACAGAGCA 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3462 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGGAATATTTCTGACAGAGCA 3521

Qy         121 GCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3522 GCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGGAG 3581

Qy         181 AGCCCAGAATGTCACTCGGCTTCACAGAAAGCAGGCTGCATACTGAGTCGAGGTCAGTGGA 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3582 AGCCCAGAATGTCACTCGGCTTCACAGAAAGCAGGCTGCATACTGAGTCGAGGTCAGTGGA 3641

Qy         241 AAAATTGAACCTGCACTCCCGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
;   APPLICANT: Jones, Karen A.
;   APPLICANT: Volkmuth, Wayne
;   APPLICANT: Walker, Michael G.
;   TITLE OF INVENTION: BONE REMODELING GENES
;   FILE REFERENCE: PB-0014 US
;   CURRENT APPLICATION NUMBER: US/09/484,970B
;   CURRENT FILING DATE: 2000-01-18
;   NUMBER OF SEQ ID NOS: 172
;   SOFTWARE: PERL Program
; SEQ ID NO 60
;   LENGTH: 13977
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
;   NAME/KEY: unsure
;   LOCATION: 11721-11761, 12294, 13969
;   OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60
```

```
Query Match           99.2%;  Score 1290;  DB 4;  Length 13977;
Best Local Similarity 99.9%;  Pred. No. 0;
Matches 1301;  Conservative 0;  Mismatches 0;  Indels 1;  Gaps 1;

Qy    1 CGACTTTCCAGCAGTTCAGAGCAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8716 CGACTTTCCAGCAGTTCAGAGCAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 8775

Qy   61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTTCTGACAGAGCA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8776 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTTCTGACAGAGCA 8835

Qy  121 GCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8836 GCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 8895

Qy  181 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGCTGAATACTGAGTGGGA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8896 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGGCTGAATACTGAGTGGGA 8955

Qy  241 AAAATTGAACCTGACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8956 AAAATTGAACCTGACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACT 9015

Qy  301 CCAGGAACTTCAAGAGGCCACGGACCCGTGGGGCGATCCTCCTCAAGTCGCCAAGCTGAGGTGAT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9016 CCAGGAACTTCAAGAGGCCACGGACCCGTGGGGCGATCCTCCTCAAGTCGCCAAGCTGAGGTGAT 9075

Qy  361 CAAGGGATCCTCGAGCGCCCGGCAGCCCGTGGGGCCGATCTCTCTCCAAGATCACCTCGA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9076 CAAGGGATCCTCGAGCGCCCGGCAGCCCGTGGGGCCGATCTCTCTCCAAGATCACCTCGA 9135

Qy  421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTCGGAAGAGAACGTGAGCCACGTCAA 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9136 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTCGGAAGAGAACGTGAGCCACGTCAA 9195

Qy  481 TGACCTTGCTCGCGCCAGCTTACCACTTGGGCCATTCAGCTTCAGCTTACCTCAGCAC 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9196 TGACCTTGCTCGCGCCAGCTTACCACTTGGGCCATTCAGCTTCAGCTTACCTCAGCAC 9255

Qy  541 TCTGGAAGACCTGAACACCAGCAGATGGAAGTTCTGCAGTGGCCGTCGAGGACCGAGTCAG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9256 TCTGGAAGACCTGAACACCAGCAGATGGAAGTTCTGCAGTGGCCGTCGAGGACCGAGTCAG 9315

Qy  601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCTCGCAGCATCCTCTTCTTCCACGTC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9316 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCTCGCAGCATCCTCTTCTTCCACGTC 9375

Qy  661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
```

```
Db 3642 AAAATTGAACCTGCACTCCCGCTGACTGCTGGCAGAGAAAATAGATGAGACCCCTTGAAAGACT 3701
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 CCAGGAACTTCAAGAGGCCACGGACCCGTGGGGCGATCCTCCTCAAGTCGCCAAGCTGAGGTGAT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3702 CCAGGAACTTCAAGAGGCCACGGACCCGTGGGGCGATCCTCCTCAAGTCGCCAAGCTGAGGTGAT 3761

Qy  361 CAAGGGATCCTCGAGCGCCCGGCAGCCCGTGGGGCCGATCTCTCTCCAAGATCACCTCGA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3762 CAAGGGATCCTCGAGCGCCCGGCAGCCCGTGGGGCCGATCTCTCTCCAAGATCACCTCGA 3821

Qy  421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTCGGAAGAGAACGTGAGCCACGTCAA 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3822 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTCGGAAGAGAACGTGAGCCACGTCAA 3881

Qy  481 TGACCTTGCTCGCGCCAGCTTACCACTTGGGCCATTCAGCTTCAGCTTACCTCAGCAC 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3882 TGACCTTGCTCGCGCCAGCTTACCACTTGGGCCATTCAGCTTCAGCTTACCTCAGCAC 3941

Qy  541 TCTGGAAGACCTGAACACCAGCAGATGGAAGTTCTGCAGTGGCCGTCGAGGACCGAGTCAG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3942 TCTGGAAGACCTGAACACCAGCAGATGGAAGTTCTGCAGTGGCCGTCGAGGACCGAGTCAG 4001

Qy  601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCTCGCAGCATCCTCTTCTTCCACGTC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4002 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCTCGCAGCATCCTCTTCTTCCACGTC 4061

Qy  661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4062 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 4121

Qy  721 CGAGACTCAAACAACTGCTGGGACCATCCCAAATGACAGAGCTCACCAGTCTTTTAGC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4122 CGAGACTCAAACAACTGCTGGGACCATCCCAAATGACAGAGCTCACCAGTCTTTTAGC 4181

Qy  781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAGACTGCA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4182 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAGACTGCA 4241

Qy  841 GAAGGGCCCTTTGCTTGGATCTCTTGAGCCTCGTCAGCTGCATGTCGCATGATGCCTGGACCAGCA 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4242 GAAGGGCCCTTTGCTTGGATCTCTTGAGCCTCGTCAGCTGCATGTCGCATGATGCCTGGACCAGCA 4301

Qy  901 CAACCTCAAGCAAAATGACCAGCCATGGACCCCATGAGCAGATTATTAATTGTTTGACCAC 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4302 CAACCTCAAGCAAAATGACCAGCCATGGACCCCATGAGCAGATTATTAATTGTTTGACCAC 4361

Qy  961 TATTTATGACCGCCTGCTGCTGTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4362 TATTTATGACCGCCTGCTGCTGTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 4421

Qy 1021 TATGTGTCTGAACTGGCTGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4422 TATGTGTCTGAACTGGCTGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 4481

Qy 1081 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAGACAAGAAGTACAG 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4482 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAGACAAGAAGTACAG 4541

Qy 1141 ATACCTTTCAGCAAGTGGCAAGTTCAACAGAGATTTTGTGACCGCGGCAGGCTGGGCCT 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4542 ATACCTTTCAGCAAGTGGCAAGTTCAACAGAGATTTTGTGACCGCGGCAGGCTGGGCCT 4601

Qy 1201 CCTTCGTCGCCAAGTTCATCGATTCCATCAAAATTCCAAAGACAGTTGGGTGAGTTGGGGG 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4602 CCTTCGTCGCCAAGTTCATCGATTCCATCAAAATTCCAAAGACAGTTGGGTGAGTTGGGGG 4661

Qy 1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
        |||||||||||||||||||||||||||||||||||||||||
Db 4662 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4702
```

RESULT 2
US-09-484-970B-60

APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-836-022A-10

Query Match 87.3%; Score 1135.2; DB 3; Length 19307;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 0; Indels 103; Gaps 0;

```
Qy      2 GACTTTCCAGCAGTTCAGGAACGAGGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 61
            ||||| |||||||||||||||||||| ||||||||| ||||||||||||||||||||||||
Db   5977 GATTTCCCAGCAGTTCAGGAACGCAGAATGAGATATACATAGGGCCTTCAAGAGGGAATTGAAA 5918

Qy     62 ACTAAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTTCTGACAGAGCAG 121
            ||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   5917 ACTAAAGAACCTGTAATCATGAGTACTCTGGAGACTGTGAGAATATTTCTGACAGAGCAG 5858

Qy    122 CCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGAGAGAGA 181
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5857 CCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGAAAGAGA 5798

Qy    182 GCCCAGAATGTCACTCGGCTTCTACGAAAGCAGAGGTGAGGAGTCAATACTGAGTGGGAA 241
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5797 GCTCAGAATGTCACTCGGCTCCTACGAAAGCAGAGGTGAAGAGGTGAAGGTGAATGGGAC 5738

Qy    242 AAATTGAACCTGCCACTCCGCTGACTGCAGAGAAAAATAGATGAGAAAGACTC 301
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5737 AAATTGAACCTGCCGCTGCTCAGCTGATTGCAGAGAAAAATAGATGAAGACTC 5678

Qy    302 CAGGAACTTCAAGAGCCACGGATGAGTCCGATGAACTGGACTGGATCTCCAAGTTGCGCC 361
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5677 CAGGAACTTCAAGAGCCACGGATGACTGGATGAGTCCGATGAACTGGATCTCCAAGTTGCGCC 5618

Qy    362 AAGGGATCCTGGCAGCCCGTGGGCGGCCGATCTCCTCCATTGACTCTCCCAAGATCACCTCGAG 421
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5617 AAGGGATCCTGGCAGCCCAGTGGGCGGCCGATCTCCTCCATTGACTCTCCCAAGATCACCTTGAA 5558

Qy    422 AAAGTCAAGGCACTTCGAGAGAGAAATTGCGCTCGAAAGAAACGTGAGCCCACGTCAAT 481
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5557 AAAGTCAAGGCACTTCGGGAGAGAAATTGCACCTCTTAAAGAAATGTCAATCGTGTCAAT 5498

Qy    482 GACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTGTATAACCTCAGCACT 541
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5497 GACCTTGCACATCAGCTGACCACACTGGGCATTCAGCTCTCCACCTTATAACCTCAGCACT 5438

Qy    542 CTGAAGACCTGAACACCAGATGGGAAGCTTCTGCAGGTGGCCTGGAGGACCGAGTCAGG 601
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5437 TTGAAGATCTGAATACCAGATGAGGCTTCTACAGGTGGCCTGTGGAGGACCGTGTCAGA 5378

Qy    602 CAGTCGCATGAGCCCACGGGACTTGGTCCTGCAGCATCCCAGCACTTTCTTTCCACGTCT 661
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5377 CAGCTGCATGAGCCCACACAGGGACTTTGGTCCTGCGATCTCCCAGCACTTCCTTTCCACTTCA 5318

Qy    662 GTCCAGGGTCCCTGGGAGAGAGCCATCTGCCCAAACAAAGTGCCCTTACTATAATCAACCAC 721
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5317 GTTCAGGGTCCCTGGGAGAGAGCCATCCACCACAAACAAAGTGCCCTACTATATATCAACCAC 5258

Qy    722 GAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTCTACCAGTCTTTAGCT 781
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5257 GAGACCCAAACCACTTGTTGGGACCACCCCCAAAATGACAGAGCTCTACCAGTCTTTAGCT 5198
```

```
Db   9376 TGTCCAGGGTCCCTGGGAGAGAGCCATCTGCCCAAACAAAGTGCCCTACTATATCAACCA 9435

Qy    721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTTACCAGTCTTTAGC 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9436 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTTACCAGTCTTTAGC 9495

Qy    781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9496 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 9555

Qy    841 GAAGGCCCTTGCTTGGATCTCCTTGAGCCTGCAGTGCCTGCAGTGATGCCTGGACCAGCA 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9556 GAAGGCCCTTTGCTTGGATCTCCTTGAGCCTGCAGTGTCAGCTGCATGCCTTGGACCAGCA 9615

Qy    901 CAACCTCAAGCAAAAATGACCAGCCATGACATCCTGCAGATATTAATTGTTGACCAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9616 CAACCTCAAGCAAAAATGACCAGCCATGACATCCTGCAGATATTAATTGTTGACCAC 9675

Qy    961 TATTTATGACCGGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCCTCTCTGCGTGA 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9676 TATTTATGACCGGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCCTCTCTGCGTGA 9735

Qy   1021 TATGTGTCTGGACTGGCTCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9736 TATGTGTCTGAACTGGCTCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 9795

Qy   1081 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAGACAAGTACAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9796 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAGACAAGTACAG 9855

Qy   1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCT-GGGCC 1199
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   9856 ATACCTTTCAAGCAAGTGGCAAGT-CAACAGGATTTTGTGACCAGCGCAGGCGTGGGGGC 9915

Qy   1200 TCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTGGGG 1259
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9916 TCCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 9975

Qy   1260 GCAGTAACATTGAGCCAAGTCCGGAGCTGCTTCCAATTTG 1301
            ||||||||||||||||||||||||||||||||||||||||
Db   9976 GCAGTAACATTGAGCCAAGTCCGGAGCTGCTTCCAATTTG 10017
```

REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match   87.3%; Score 1135.2; DB 3; Length 19307;
Best Local Similarity   92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```
Qy      2 GACTTTCAGCAGTTCAGAAGCAGAAGCGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 61
          |||||| |||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   5977 GATTTCCAGCAGTTCAGAAGCAGAAGCAGATGATATACATAGGGCCTTCAAGAGGGAATTGAAA 5918

Qy     62 ACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAG 121
          |||||||||| |||||||||| ||||||||||||||||| |||||||||||||||||||
Db   5917 ACTAAAGAACCTGTAATCATGAGTACTCTGGAGACTGTACCAGTATTCTGACAGAGCAG 5858

Qy    122 CCTTTGGAAGACTAGAGAAACTCACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGAGA 181
          |||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   5857 CCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAAGAGAAGA 5798

Qy    182 GCCCAGAAGTGCACTCGGCTTCACGGCTTGGCCAGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAA 241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5797 GCTCAGAATGTCACTCGGCTCCTCGGCCTCTCGGCCAGAAAGCAGGCTGAAGAGGTCAACGCTGAATGGGAC 5738

Qy    242 AAATTGAACCTGCACTCCGTTGACTGGCCAGAGAAAATAGATGAGACCTTGAAAGACTC 301
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   5737 AAATTGAACCTGCGCTCCAGCTGATTGGCCAGAGAGAAAATAGATGAAGAGCTCTTGAAAGACTC 5678

Qy    302 CAGGAACTTCAAGAGGCACGGATGAGCTGCGCCAAGCTGAGTGATC 361
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5677 CAGGAACTTCAGGAGGCACGGATGAGCTGCGCCAAGCTGAGGTGATC 5618

Qy    362 AAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCGAAGATCACCTCGAG 421
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   5617 AAGGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTGACTCTCTCGAAGATCACCTTGAA 5558

Qy    422 AAAGTCAAGGCACTTCAGAGACGTGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAAT 481
          ||||| |||||| |||||||||| |||||||||||| ||||||||||||||||||||||||||||
Db   5557 AAAGTCAAGGACTTCGGGGAGAAATTGCACCTCTTAAGAGAAGAATGTCAATCGTGTGTCAAT 5498

Qy    482 GACCTTGCTCGCCAGCTTACCACTTTGGTCCAGACACTTGGCATTCAGCTCTCACCGTATAACCTCAGCACT 541
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   5497 GACCTTGCACATCAGCTGACCACCACTGGCATTCAGCTCTCACCACCTTATAACCTCAGCACT 5438

Qy    542 CTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGGAGACCGAGTCAGG 601
          ||||||||||| |||||||||||||||||||||||||||||||| ||||||||||||||
Db   5437 TTGGAAGATCGAATACCAGATGGAGGCTTGTGGGGAGACCGAGTCGAGA 5378

Qy    602 CAGCTGCATGGAGCCCACAGGGACTTTGGTCCAGACCATCCCAAATGACGAGAGCTCTTAGCT 661
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   5377 CAGCTGCATGGAGCCCACCACAGGGACTTTGGTCCTGACCATCCCAAAATGACACACCCAAAATGACAGAGCTCTTTAGCT 5318

Qy    662 GTCCAGGTCCCTGGGAGAGCCCATCTCGCCAAACAAGTGCCTACTATATCAACCAC 721
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   5317 GTTCAGGTCCCCTGGGAGAGCCCATCTCACCAAACAAAGTGCCCTACTATATCAACCAC 5258

Qy    722 GAGACTCAAAACACTTGCTGGGACCATCCCAAAATGACGAGAGCTCTACCAGTCTTTAGCT 781
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   5257 GAGACCCAAAACCACTTGTTGGGACACCCCAAAATGACACAGAGCTCTACCAGTCTTTAGCT 5198

Qy    782 GACCTGAATAATGTCAGATTCAGAATTCCGCGGTTATAGGACGTTATAGGACTCCCGGGTATAGGACTCCAG 841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5197 GACCTGAATAATGTCAGGTTCAGAATTCCGCGGTTATAGGACTCCCGGGTATAGGACTCCAG 5138
```

RESULT 4
US-09-427-048A-10/c
Sequence 10, Application US/09427048A
Patent No. 6203975
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
Wilson, James M.
Fisher, Krishna J.
Chen, Shu-Jen
Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
Methods of Use Thereof
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215

Qy   2 GACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGATTGAAA 61
        |||||| ||||||| |||| ||| | |||||||| | ||| |||||| ||
Db 3506 GATGTTCCAGCCTTACAGCTCCAGTATGACCATTGTAAGGCCCTGAGACGGGAGTTAAAG 3565

Qy   62 ACTAAAGAACCTGTAATCACATGAGTACTCTGAGACTGTACGACTGTTGTTGACCACT 961
        ||| ||| |||| ||||||||||
Db 3566 GAGAAGAATATTCTGTCCTGAATGCTGTCGACCAGGCCCGAGTTTTCTTGGCTGATCAG 3625

Qy   122 CCTTTGGAAGGACTAGAGAAACT------CTACCAGGAGCCCAGAGAGCTGCCTCCT 1021
        ||||| ||||||
Db 3626 CCAATTGAGGGCCCCTGAGAGCCAAGAAGAAACCTACAATCAATCAGAATCAGAATTAACTCCT 3685

Qy   173 GAGGAGAGAGCCCAGATGTCACTCGGCTTCTACCGAAAGCAGGCTGAGGAGGTCAATACT 232
        ||||||
Db 3686 GAGGAGGAGCGAGCCCAAAGAGATTCCAAAGCCATGCCGACTGGCAGAAGCAAGCATTG 3745

Qy   233 GAGTGGGAAAAAATTGACCTGCCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTT 292
        ||||||
Db 3746 AAATGGGAAAGTCTAAATGCTGTAACTAGCACATTGCAAAAGCAAGTGGACATTG 3805

Qy   293 GAAAGACTTCAGGAACTTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGCT 352
        ||| ||
Db 3806 GAGAAACTCAGAGACCTGCAGGGAGCTATGATGACGACTGGACGCGTGACATGAAGAGGGCA 3865

Qy   353 GAGGTGATCAAGGGATCCTGGCCAGCCGTGTGGCGATCTCCTCATTGACTCTCTCCAAGAT 412
        ||| |||||| |||
Db 3866 GAGTCCGTGCGGAATGGCTGGAAGCCCGTGGGGAGGACTTACTCATTCATGACTCGCTGCAGGAT 3925

Qy   413 CACCTCGAAAGTCAAGGCACTTCGAGGGAGAAATTGCGCCTCTGAAGAGAGAACGTGAGC 472
        ||| ||
Db 3926 CACATTGAAAAAATCATGGCACTTAGAGAAGAAATTGCACCAATCAACTTTAAAGTTAAA 3985

Qy   473 CACGTCAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAAC 532
        ||| ||
Db 3986 ACGGTGAATGATTTATCCGATCCAGCTGTCTCCACTTGACCTGCCGCCTGGCCGTCGAGGAC 4045

Qy   533 CTCAGCACTCTCGAAGATCTCTGGAAGCCAGAGCTCCGAAGTCTTCGCAGGTGGCCGTCGAGGAC 592
        ||| ||
Db 4046 TCTCGCCAGCTAGATGACCTTAATATGCGATGGAAACTTTACAGGTTTCTGTGGATGAT 4105

Qy   593 CGAGTCAGGCAGCTGCATGAGAGCCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTT 652
        ||| ||
Db 4106 CGCCTTAAACAGCTTCAGGAGCCCCAGAGATTTTGGACCATCCCTCCAGCATTTTCTC 4165

Qy   653 TCCACGTCGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 712
        ||| ||
Db 4166 TCTACGTCCAGTCCAGCTGCCGTGGCAAAGATCCATTTCACATAATAAAGTGCCCTATTAC 4225

Qy   713 ATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAG 772
        ||| ||
Db 4226 ATCAACCATCAAACACAGACCACCTGTTGGGACCATCCTAAAATGGCCGAACTCTTTCAA 4285

Qy   773 TCTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTTCTTCTGCCATGAAACTCCGA 832
        ||| ||
Db 4286 TCCCTTGCTGACCTGAATAATGTACGTTTTTCTGCCTACCTCACCGTACACAGCAATCAAAATCCGA 4345

Qy   833 AGACTGCAGAAGCGCCCTTGCTGGATCTCTTCTTGAGCCTGTCAGCTGCATGTGATGGCCTTG 892
        ||| ||
Db 4346 AGACTACAAAAAGCACTATGTTTGGATCTCTTAGAGTTGAGTACAACAAATGAAATTTTC 4405

Qy   893 GACCAGCACAACCTCAGCAAGGAAATGAACAAAATGAACCAGCTCCTCCAGTGTTCCAGATGTCATCAACTGT 952
        ||| ||
Db 4406 AAACAGCACAAGTGACCAAAATGAACCAAAATGAACCAGCTCCTCCAGTGTTCCAGATGTCATCAACTGT 4465

Qy   953 TTGACCACTATTATGACCGGCTGCTGAAGTGTTTATGATACGGAGCAACAACATTTGGTCAACGTCCCTCTC 1012
        ||| ||
Db 4466 CTGACAACAACTTATGATGGACTTGAGCAAATGCATAAGACCTGGTCAACGTTCCACTC 4525

Qy   1013 TGCCTGGATATGTGTCTGAACTGGCTTGCTGAAGTTTATGATACGGGACGAACAGGGAGG 1072
        ||| ||
Db 4526 TGTGTTGATATGTGTCTCAATTGGGTTGCTCAATGGTGCTATGTCTTATGACACGGGTCGAACTGGAAAA 4585

Qy   842 AAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGGCCTTGGACCAGCAC 901
        ||| ||
Db 5137 AAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGTGCTCAGCTGCATGTGATGGCCTTGGACCAGCAC 5078

Qy   902 AACCTCAAGCAAAATGACCAGCCATGGAATCTGCAGATTAATTGTTGACCACT 961
        ||| ||
Db 5077 AACCTCAAGCAAAATGACCAGCCATGGAATCCTGCAGATAATTAACTGTTGACTACA 5018

Qy   962 ATTTATGACGGCTCGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTTCTGCTGTGGAT 1021
        ||| ||
Db 5017 ATTTATGATCGTCGGAGCAAGAGCACAACAATCTGTCAACGTCCCTCTCTGTGTGGAT 4958

Qy   1022 ATGTGCTGCTGGCTGCTCGAATGTTTATGATACGGGACGAACAGGGAGGGATCCGTGTC 1081
        ||| ||
Db 4957 ATGTGTCTCAACTGGGTTCTCAATGTTTATGATACGGGACGAACAGGGACGGGATCCGTGTC 4898

Qy   1082 CTGTCTTTTAAAACTGGCATCATTCCCTGGTAAAGCTGTAAAAGCAAGTACAGA 1141
        ||| ||
Db 4897 CTGTCTTTTAAAACTGGCATCATTTCTCGTGTAAAGCTCCACACTTGCAAGTACAGA 4838

Qy   1142 TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGCACCAGCGCAGGCTGGGCCTC 1201
        ||| ||
Db 4837 TACCTTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTGTGACCAGGCTAGGCTGGTCTT 4778

Qy   1202 CTTCTGACGATTCTATCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGGC 1261
        ||| ||
Db 4777 CTTCTGCATGATGAATTCTATTCAAATCCCAAGACCAGTGGGTGAAGTTGCTTCCTTTGGGGGC 4718

Qy   1262 AGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
        |||||||
Db 4717 AGTAACATTGAGCCGAGTGTCAGGAGCTGCTTCCAATTTG 4678

RESULT 5
US-09-091-501B-7
Sequence 7, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 6045
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
NAME/KEY: misc_feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match   42.6%;   Score 553.8;   DB 4;   Length 6045;
Best Local Similarity   64.7%;   Pred. No. 1.1e-169;
Matches 844;   Conservative   0;   Mismatches 452;   Indels 9;   Gaps 1;

Qy   1073 ATCCGTGTCCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGAC 1132
          ||||| |||||| ||| ||||| |||| ||||| ||| | |||| ||||||||||||||
Db   4586 ATTAGAGTGCAGAGTCTGAAGATTGGATTAAATGTCTCTCCAAAGGTCTCTTGGAAGAA 4645

Qy   1133 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTGGCCAAGTTCAACAGGATTTTGTGACCAGCGCAGG 1192
          |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4646 AAATACAGATATCTCTTTAAGGAAGTGTCGGGGCCGACAGAAATGTGTGACCAGTAGCAGCT 4705

Qy   1193 CTGGGCCTCCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCC 1252
          |||||| ||||||||||||||||| || |||||||||||||||||||||||||||||||
Db   4706 CTGGGCCTGTTACTTCATGATGCCATCCATCCAGATCCCCGGCCGGTAGGTAGGTAGCAGCT 4765

Qy   1253 TTTGGGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAA 1297
          ||||||| ||||||||||||||||||||||||||||||||||||
Db   4766 TTTGGAGGCAGTAATATTGAGCCTAGTGTTCGCAGCTGCTTCCAA 4810

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
;   APPLICANT: Tingley, Jonathon M
;   APPLICANT: Davies, Kay E
;   TITLE OF INVENTION: Utrophin gene expression
;   FILE REFERENCE: 620-42
;   CURRENT APPLICATION NUMBER: US/09/091,501B
;   CURRENT FILING DATE: 1998-06-18
;   PRIOR APPLICATION NUMBER: PCT/GB96/03156
;   PRIOR FILING DATE: 1996-12-19
;   PRIOR APPLICATION NUMBER: GB 9525962.8
;   PRIOR FILING DATE: 1995-12-19
;   PRIOR APPLICATION NUMBER: GB 9615797.9
;   PRIOR FILING DATE: 1996-07-26
;   PRIOR APPLICATION NUMBER: GB 9622174.2
;   PRIOR FILING DATE: 1996-10-24
;   NUMBER OF SEQ ID NOS: 15
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 9
;   LENGTH: 10320
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (11)..(10312)
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Full length
;   OTHER INFORMATION: utrophin construct
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (724)...(758)
;   OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match   42.6%; Score 553.8; DB 4; Length 10320;
Best Local Similarity   64.7%; Pred. No. 1.5e-169;
Matches 844; Conservative 0; Mismatches 452; Indels 9; Gaps 1;

Qy   2 GACTTTCCAGCAGTTCAGAGAACGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 61
          ||||||||||||||||| |||| ||||| ||||| |||| | ||||||||||||||||||||
Db   7781 GATGTTCCAGCCTTACAGCTTCAGACTGTACTCTTGACAGACTGTCTTGATGATGCCTTGAAGG 7840

Qy   62 ACTAAGAACCTGTAATCATGATGATTCTCTTGACCTGTCCCTGAATGCTGTCGACCCAGGCGTCCTGAATGCTG 121
          |||| ||||| ||||| | |||| ||| || | || ||| |||||||||||||| |||||||
Db   7841 GAGAAAGAATAATTCTGTCCCTGGAATGCTGTCGACCCGAGTTTTCTTGGCTGATCAG 7900

Qy   122 CCTTTGGAAGGAGCTAGAGAAACT-------CTACCAGGAGGCCCAGAGAGTGCCTCCT 172
          |||| |||||||||| |||||||| |||| ||||||| |||||||||||||
Db   7901 CCAATTGAAGGCCCTGAAGAGCCCAGAGAAGCCAGAAGCAGCTGAGGAGGTCAATACT 7960

Qy   173 CAGGAGGAGAGCCCCAGAAATGCCAAAGATTGCCAAAGCCATGCGCAAAGCAGTCTTCTTGAAGTCAAAGAA 232
          |||||||||||| | ||||| ||| |||||| | |||||||| ||| ||||||||||||| ||| ||||
Db   7961 GAGGAGGAGAGCCCAAGATTTCTGAAGTCAAGAGAACAGTCTTCTTGAAGTCAAAGAA 8020

Qy   233 GAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCGAGAGAAAAATAGATGAGACCCTT 292
          ||| | ||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   8021 AAATGGGAAGGTCTAAATGCTGAAATGTCTCTCCAAAGGTGGCAAGGCATTG 8080

Qy   293 GAAAGACTTCCAGGAACTTCAAGCAACTGGAGATGGCCACGACGATGAGCTGGACCTGGACCTGGACGCTGGACGTGCGCGCAAGCT 352
          ||| |||||||||| ||||| ||| || ||||| ||||| || |||||||||||||||||||||
Db   8081 GAGAAACTTCAGAGACCTCAGGAGGTGCATGACGATGAGCTGACATGAAGGAGGCA 8140

Qy   353 GAGTGATCAAGGGATCCTGGCACGCCGTGGGCGATCTCCTTCATTGACTCTTCCCAAGAT 412
          ||| ||||||||||||| |||| |||||||||||||||| | ||||||||||||| |||
Db   8141 GAGTCCGTGCCGAAGTGCCTGGGAGATGAAAATTGAGTCCTGGGGAGGACTTACTCATTGACTCGCTCGAGGAT 8200

Qy   413 CACCTGCGAGAAAGTCAAGCACTCAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGC 472
          |||| ||||| |||| | ||||| || | ||| || || || |||||||| ||||||
Db   8201 CACATTGAAAAAATCATGGCATTTGACCAATCACTTTAAAGTTAAA 8260

Qy   473 CACGTTCAATGACCTTGCTCGCGCCAGCTTACCACTTTGGGCATTCAGTCTCTCACCGTATAAC 532
          ||||| ||| || | || || | || | |||| |||||||||| ||||| || |||||
Db   8261 ACGGTTGAATGATTTATCCAGTCAGTCGTGTCTCCACTTGACCTGCATCCCTCTCTAAAGATG 8320

Qy   533 CTCAGGACACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGAC 592
          ||| || ||||||||||||||||||||||| ||||| ||||||| ||| |||||| |||
Db   8321 TCTCGCCAGCTAGATGACCTTAATATGCGATGGAAACTTTTACAGGTTTCTGTGGATGAT 8380

Qy   593 CGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCGTCCAGCATCCAGCACTTTCTT 652
          ||||||||||| ||| |||| ||||| ||||| |||| |||||||| || || |||||||
Db   8381 CGCCTTAAACAGTCTCAGGAAGCCCACCACCAGATTTTGGACCATTTCCTCTCAGCATTTTTCTC 8440

Qy   653 TCCACGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTAT 712
          ||||| || |||||||| |||||| ||||| ||||| | ||||| ||||||||||| |
Db   8441 TCTACGTCAGTCCAGCTGCCGTGCAAGATCCATTCACATATAAAGTGCCCATTAC 8500

Qy   713 ATCAACCACGAGACTCAAACAACTTGCTGGGACAACTACACTGCCAAAATGACAGAGCTCACCAG 772
          |||||||||| ||||||||| |||||| ||||| ||||| | ||||||||| |||||||| |||
Db   8501 ATCACCATCAAACACAGACCACGTGTGGGACCATCCTAAAAATGACCGAACTCTTTCAA 8560

Qy   773 TCTTTAGCTGACCTGAATAATGTCAGATTCAGATTTGTTTTCTGCCTACCGTACAGACTGCCATGAAAACTCCGA 832
          ||| | ||||||||| | |||| | | ||||| ||| | || ||||| ||| ||| ||||| |||||||
Db   8561 TCCCTTGCTGACCTGAATAATGTACGTTTTTGTCTTGGGATCCTTGAGCTCTTGAGCTC 8620

Qy   833 AGACTGCAGAAGGCCCTTGCTTGATCCCTTGTGCTTGAGCAGTGATGATGATGCGTTG 892
          |||||| || | | ||||||||||| || || |||| ||||||||| || ||||||
Db   8621 AGACTACAAAAAGCACTATGTTTGGATCTCTAGAGTTGAGTCAACAAATGAAATTTTC 8680

Qy   893 GACCAGCACAACCTCAAGCAAATGACCAGCCCATGCAAAATGAACCAGTGTCATCATCATCAACTGT 952
          ||||| || ||||||||||| | ||||||||||| || |||||||||| |||||||||||||||
Db   8681 AAACAGCACAAGTGAACCAAATGACCAGCTCCTCAGTGTTCCAGATGTCATCAACTGT 8740

Qy   953 TTGACCACTATTATGACCGCCTGGAGCAAGAGCACAACAAATTGGTCAACATTGGTTCAACGTCCCCTC 1012
          |||||| |||||| ||||||| ||||| || | | ||||| ||||| ||||| |||| |||||||
Db   8741 CTGACAACAACTTATGATGGACTTGAGCAAATGCATAGGCCATGGCTTGGACTTGGCTCTGGGCTCACTC 8800

Qy   1013 TGCGTGGATATGTGTCTGAACTGCTGCTCGTGAATGTTTATGATACGGACGAACAGGAGG 1072
          ||| |||| |||||| | ||||||| ||||||||||| |||| |||| || ||||||||
Db   8801 TGTGTTGATATGTGTCTCAATTGGTTGCTCAATTGGTCTCATGACAGGAAA 8860

Qy   1073 ATCCGTGTCCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGAC 1132
          |||||| || | |||| | | |||| ||| | || | |||| ||||||||||||||||
Db   8861 ATTAGAGTGCAGAGTCTGAAGATTGGATTAAATGTCTCTCCAAAGGTCTCTTGGAAGAA 8920

Qy   1133 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTGGCCAAGTTCAACAGGATTTTGTGACCAGCGCAGG 1192
          ||||||||||| || |||||||||||||| ||||||| ||||||| ||||| |||||||||||||||
Db   8921 AAATACAGATATCTCTTTAAGGAAGTTCGGGGCCGACAGAAATGTGACCAGTAGCAGCAG 8980

Qy   1193 CTGGGCCTCCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCC 1252
          |||||||| || |||||| || || |||||| |||| |||| ||||||||||| |||
Db   8981 CTGGGCCTGTTACTTCATGATCATGAGCCATCCATCCAGATCCCCGGCCGCAGTAGCAGCT 9040

Qy   1253 TTTGGGGCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAA 1297
          |||||| ||||| ||||||||||||||||| |||||||||||||
Db   9041 TTTGGAGGCAGTAATATTGAGCCTAGTGTTCGCAGCTGCTTCCAA 9085

RESULT 7
US-09-976-594-93
; Sequence 93, Application US/09976594
; Patent No. 6673549
GENERAL INFORMATION:
; APPLICANT: Buchbinder, Michael
; APPLICANT: Furness, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
US-09-976-594-93

; SEQ ID NO 93
; LENGTH: 3915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 290344.1
US-09-976-594-93

Query Match   41.5%;   Score 540.2;   DB 4;   Length 3915;
Best Local Similarity 64.3%;   Pred. No. 2.3e-165;
Matches 827;   Conservative   0;   Mismatches 458;   Indels   2;   Gaps   1;

Qy   14 GTTCAGAAGCAGAACGATGCTACATAGGGCCTTCAGAGGGGAATTGAAAACTAAAGAACCT   73
          ||| | ||| || | | ||| || ||||||| ||||| ||||| | |||| |||
Db   722 GTGCAACAGGAGAAGGAGACACATGCCGCCTTTATGGAAGAAGTCAAGTCTCGGGGCCCC   781

Qy   74 GTAATCATGAGTACTCTTGAGACTGTACGAGATATTCTGACAGAGCAGCCTTTGGAAGGA   133
          | | ||||| | | |||| || | || | || ||| || | |||||| |||||
Db   782 TACATCTATTCTGTGTCTGGAGTCCAGCTGAGTCCTCCTGTCCAGCACCCATTTGAGGAG   841

Qy   134 CTAGAGAAACTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAGAGGAGCCCAGAATGTC   193
          || ||| || | || |||| ||| | ||| | | |||| || | || ||||| |||
Db   842 TTAGAGGAGCCTCATTCTGAGAGCAGAGCAGCAAAGATACCTCCCGAAACAGCCGATCCAGAAATCTC   901

Qy   194 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAACCTG   253
          | | | | | | || |||| |||| ||| | ||| |||||||| |||||| ||
Db   902 AGCCGCTTTGTATGGAAGCAGGCAGGCCGACGGTGGACGGTGGCCAGTGAACTGTGGGAGAGTTGACAGCC   961

Qy   254 CACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAA   313
          | | | | | | ||| | | || || | ||| | ||||| || ||||| || ||
Db   962 CGCTGTGTGGACCAGCACGCACCGTCACATTGAGCCGACTTCTGGAGCTCTTGGAGATTCA-   1020

Qy   314 GAGGCCACGGATGAGTCGTGGACCTGCTGCTACCTCAAGCTCGCGCCAAGCTCGAGGTGATCAAGGGATCCTGG   373
          || | | | | | | ||| | | || || |||| | |||| | | || ||| ||
Db   1021 -GGGGCATGGAGGAACTAAGCACTACTCTGAGCCAAGCTGAGGGAGTCCCAGCCCACTTGG   1079

Qy   374 CAGCCCGTGGGCGCCATCTCCTCATTGACTCTCCAGATCACCTCCAGAAAGTCAAGGCA   433
          | | | | ||| | ||| || | ||| | | | || | ||| | ||| |
Db   1080 GAGCCCATTGGGCGATCTCTTCTTCATTGATTCACCCCGAGCACATCCAGGCTATTAAGCTG   1139

Qy   434 CTTCGAGAGAAAATTGCGCCTCTGAAAGAGAACGTGAGCCCACGTCAATGACCTTGCTCGC   493
          | || ||| || ||||| || || || || | | || | | | | | || | | |
Db   1140 TTCAAAGAAGAATTCTCCCCATGAAAGAAGATGAGGTAAAGTTGTGAAATGATCTGCCCAC   1199

Qy   494 CAGCTTACCACTTTGGGCATTCAGCTCACCGTATAACCTCGAAGACCTG   553
          | || | || ||| ||| | | | | | || | ||| ||
Db   1200 CAACTTGCCATTTCTGATGTGCACTTGTCAATGGAGAATTCCCAGGCCCCTGGAACAGATC   1259

Qy   554 AACACCAGATGGAAGCTTCTGCGAGGTGCCGTCGAGGCAGCTCGCATGAA   613
          || || | ||| ||| || | || || | | ||| ||| | | ||
Db   1260 AACTTCCGATGGAAGAACACTACAGGCGCGTCAGTTGATGGAGAGGCTTAAGCAGCTCCAGGAT   1319

Qy   614 GCCCAGAGGGACTTGGTTGGGCCTCCAGCATCCAGCAC   673
          || || | | ||| | || ||||| | | | | ||
Db   1320 GCCCACCACGGGACTTTGGGCCTGGCCTGGGAGTCACAGCACTTTCTCCTCCTCCTGTCCTGTCCAGGTTCCC   1379

Qy   674 TGGGAGGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCACGAGACTCAAACA   733
          |||| | | |||||||||| ||| | |||| ||||| | || | ||| |||||
Db   1380 TGGGAAAGAGCAATTTCACCCAATAAAGTTCCCTACTACATACCACCACCAGCGGCTCAGACC   1439

Qy   734 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT   793
          | | ||| | | | | ||| | | | || || |||| | | | || | | |
Db   1440 ACATGCTGGGACCATCCCAAGATGACAGAGAGTATATCCCTAGCCTAGCCTAGCTGGATCTGACAAC   1499

Qy   794 GTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAGAACTGCAGAGACTGGAGGGCCTTTGC   853
          | | | || |||| || || || | | | | | || || | | | || | || |
Db   1500 ATTAAGTTCTCAGCTTATCCGCACTGCCATGAAACTTCCCGCAGTCCCAGAAAGCCCTGCGC   1559

Qy   854 TTCGATCTCTTGAGCCTGTCAGCTGCTGTGAGCCTTGGCACGTGTGATGAGGTAGAGGTCAAGCAA   913
          ||| | || ||| | | | | | | || |||| | | | | ||| | | ||| ||||||
Db   1560 TTCGGACCTGGGTAACTTTAACCACCAGCCCTGGAAATCTTCAATGAGCATGAGATCTGCAGGCC   1619

Qy   914 AATGACCAGCCCATGAGATATCCTGCAGATTATTAATTGTTTGACCACTATTTATGACCGC   973
          || | | | | | ||| | | | ||||| ||| | ||| | || |||||
Db   1620 AGTGAGCACGTGATGGATGTGTGTAGAGGTCATTCACTGCGCTGACATGTGTCTGAACGT   1679

Qy   974 CTCGGAGCAAGAGCACACAACAATTTGGTCAACGTCCCTCTTGCGCTGGATGTGTGTCTGAAC   1033
          || || | || | | | | | ||| ||| | | || || | | || | | | ||
Db   1680 TTGGAGGAGGAAAGAGGCATCCTGGTCAACGTGCCACTCGTGTGACATGAGCCTCAAT   1739

Qy   1034 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCGTGTCTTTTAAA   1093
          || | | | || | || || | | | | | | | || || | | | ||| |||
Db   1740 TGGCTCCTCAAATGTTTTTTGATAGTAGTGGTCGCAGCGCAAAGATGGCGGGGCATTGTCTTTTTAAG   1799

Qy   1094 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACACAGTACAGACAAGTACAGATACCCTTTTCAAG   1153
          || || | | || | ||||| | | | |||| | || | || | | | | || | || ||||
Db   1800 ACTGGCATTGCTCATGCTTGTGTGGCACGGAAGTGAAGGAAAAACTTCAGTACCTCTTCAGC   1859

Qy   1154 CAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGGCCTTGGGCCTCCTTCTGCATGAT   1213
          |||||||| ||| | | | || | || | | | ||| | ||| ||| | |||
Db   1860 CAAGTGGCCAACTCAGGCAGCAGCCAGTGTGACCAGCGCCAGCCCCACCTTGGTGTGTCCTTCATGAG   1919

Qy   1214 TCTATCCAAATTCCAAGACAGCAGTTGGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAG   1273
          || | |||| | | | ||| | |||| | | | | || |||| | | | |||||
Db   1920 GCCATTCAGTTGCCCGTCAGCTGGGTGAAGTTGGGGTGAAGTTGGGGTGGGGGGCAGCAATGTGGAG   1979

Qy   1274 CCAAGTGTCCGGAGCTGCTTCCAATT   1300
          || | ||| ||| |||||| | |||
Db   1980 CCCAGTGTTCCGTAGTTGCTTCCGTTTT   2006

RESULT 8
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22311-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313

```
;   FILING DATE:
;     APPLICATION NUMBER: EP 91 114 300.6
;   FILING DATE:        26-AUG-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BENT, Stephen A.
;     REGISTRATION NUMBER:  29,768
;     REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:   (703)836-9300
;     TELEFAX:     (703)683-4109
;     TELEX:       899149
;   INFORMATION FOR SEQ ID NO: 14:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 7218 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     CLONE: pT2gpt-Pls
US-08-232-463-14

Query Match             3.8%;  Score 50;  DB 1;  Length 7218;
Best Local Similarity   4.2%;  Pred. No. 2.7e-05;
Matches   14;  Conservative  191;  Mismatches  131;  Indels  0;  Gaps  0;

Qy      18 AGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAA  77
              :::::.: .::: .:::        ::::::::  :::::: :::::: .::: :::
Db    1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1322

Qy      78 TCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAG 137
                 : :::: .:: ::::       ::::::::::: :::::::::::::: ::
Db    1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262

Qy     138 AGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGCCCAGAATGTCACTC 197
              ::: ::  : :::::   ::::      :::::::::: : : ::: :: :
Db    1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202

Qy     198 GGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACT 257
              ::  .: :::. ::: :: :      ::: ::  :: ::::::: :
Db    1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1142

Qy     258 CCGCTGACTCGCAGAGAAAAATAGATGAGACCCTTGAAAGACTTCCAGGAACTTCAAGAGG 317
              :: :::: ::::::::::: :::: :::::: ::::::::::::::: :: ::
Db    1141 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1082

Qy     318 CCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTG 353
              ::::: :   :::: :::  ::::::: ::
Db    1081 RRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCCCTGACCCTG 1046


RESULT 9
US-09-668-313A-10
;  Sequence 10, Application US/09668313A
;  Patent No. 6503756
;  GENERAL INFORMATION:
;    APPLICANT:  Brett P. Monia
;    APPLICANT:  Susan M. Freier
;    APPLICANT:  Jacqueline Wyatt
;    TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
;    FILE REFERENCE: RTS-0127
;    CURRENT APPLICATION NUMBER: US/09/668,313A
;    CURRENT FILING DATE:  2000-09-22
;    NUMBER OF SEQ ID NOS: 247
;  SEQ ID NO 10
;    LENGTH: 2574
;    TYPE: DNA
;    ORGANISM: Mus musculus
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (218)...(1891)
US-09-668-313A-10

Query Match             3.4%;  Score 44.2;  DB 4;  Length 2574;
```

```
Best Local Similarity   54.7%;   Pred. No. 0.0011;
Matches   88;  Conservative  0;  Mismatches  73;  Indels  0;  Gaps  0;

Qy     674 TGGGAGAGAGCCATCTCGCCAAACAACAAGTGCCCTACTATATCAACCACGAGACTCAAACA 733
              |||||||||| |||| | ||||||||||| || |||||| ||| ||||||| |||||
Db    1733 TGGGGAGAAGCTTACACAGCAGATGGAATCAACTACTTCATCAACCACGTGACACAGACC 1792

Qy     734 ACTTGCTGGACCATCCAAAATGACAGAGCTTCACCAGTGTTAGCTGACCTGAATAAT 793
              |||| |||||| |||||  ||| ||| ||| || ||||| ||| ||||||||||
Db    1793 ACGTCCTGGATCACCCCGTGATGAGCGCCCTGATGAGGCGTCCTGTCCTGTGCGAGGAGAGTGAA 1852

Qy     794 GTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAG 834
              ||| |||| | ||| ||| |||| ||||| || |||| ||
Db    1853 GAGGACTGTCCCAGAGAGCTAACAGACCCGAAAAGCTGATG 1893


RESULT 10
US-09-620-312D-69
;  Sequence 69, Application US/09620312D
;  Patent No. 6569662
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Y. Tom
;    APPLICANT:  Liu, Chenghua
;    APPLICANT:  Asundi, Vinod
;    APPLICANT:  Zhang, Jie
;    APPLICANT:  Ren, Feiyan
;    APPLICANT:  Chen, Rui-hong
;    APPLICANT:  Zhao, Qing A.
;    APPLICANT:  Wehrman, Tom
;    APPLICANT:  Xue, Aidong J.
;    APPLICANT:  Yang, Yonghong
;    APPLICANT:  Wang, Jian-Rui
;    APPLICANT:  Zhou, Ping
;    APPLICANT:  Ma, Yunqing
;    APPLICANT:  Wang, Dunrui
;    APPLICANT:  Wang, Zhiwei
;    APPLICANT:  John Tillinghast
;    APPLICANT:  Drmanac, Radoje T.
;    TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;    TITLE OF INVENTION:  Polypeptides
;    FILE REFERENCE: 784CIP2B
;    CURRENT APPLICATION NUMBER: US/09/620,312D
;    CURRENT FILING DATE:  2000-07-19
;    PRIOR APPLICATION NUMBER: 09/552,317
;    PRIOR FILING DATE: 2000-04-25
;    PRIOR APPLICATION NUMBER: 09/488,725
;    PRIOR FILING DATE: 2000-01-21
;    NUMBER OF SEQ ID NOS: 1105
;    SOFTWARE: pt_FL_genes Version 1.0
;  SEQ ID NO 69
;    LENGTH: 1690
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (128)..(1522)
US-09-620-312D-69

Query Match             3.3%;   Score 42.8;   DB 4;  Length 1690;
Best Local Similarity   47.9%;  Pred. No. 0.0023;
Matches  156;  Conservative  0;  Mismatches  167;  Indels  3;  Gaps  1;

Qy     301 CCAGGAACTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGGAGGTGAT 360
              ||||| | |||||| || |||| ||||| |  |||||||||||||||||
Db      55 CAAGGAGTTGCACCAGGTGGCGCACGACCTGGACGAGCTGGACCTGGTTCAGGAGCG 114

Qy     361 CAAGGGA--TCCTGGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCCAAGATCACCT 417
              ||||||   ||| || || |||||||| ||| || || || |||| ||||  |||||
Db     115 GCTGCCACTGCCATGCAGCAGAGCGCAGAGAGCGCAACAGGTTTGCAGGCGGTTCAGCAGCACAT 174

Qy     418 CGGAGAAAGTCAAGGCACTTCGAGGAGAGAAATTGGGCCTGAAAGAGAGACGTGAGCCACGT 477
              ||||| | ||||| ||| |||| || || ||||||||| |||||||| || || |||| |||
Db     175 CAAAAGAACCAGGCCTGCCTGCGGGCGGGATCCAGGGGATCAGCGGTGGCATGGGCCGCGCCTGGAGGAGGT 234
```

```
>y    478  CAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAG  537
             ||||||||||||| || | ||| ||||||
>b    235  GCTGGAGCGGCGCGGGGCGCGCTGGCGTTGCGTGGGAGGCTGGCGCAGAGGCAGTGCGCCG  294

>y    538  CACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCTCGAGGACCGAGT   597
             |||| | |||||
>b    295  GGGCCTGGAGCAGCTGCGAGAGCGGCCCTGGCCCGGACTGCCGGGGAGGGTGCCGACGGCA  354

>y    598  CAGGCAGCTGCATGAAGCCCACAGGG  623
             ||| |||||  ||| ||
>b    355  GCAGGTGCTGGACGCGCCGCCTTCCAGG  380

RESULT 11
US-09-368-590-1
Sequence 1, Application US/09368590
Patent No. 6187563
GENERAL INFORMATION:
   APPLICANT: Solimena, Michele
   TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
   TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
   FILE REFERENCE: 101918-200 (OCR-941)
   CURRENT APPLICATION NUMBER: US/09/368,590
   CURRENT FILING DATE: 1999-08-04
   EARLIER APPLICATION NUMBER: 60/095,657
   EARLIER FILING DATE: 1998-08-07
   NUMBER OF SEQ ID NOS: 8
   SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 1
   LENGTH: 7812
   TYPE: DNA
   ORGANISM: Human
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (1)...(6879)
   NAME/KEY: unsure
   LOCATION: (100)...(102)
   NAME/KEY: unsure
   LOCATION: (1021)...(1023)
   NAME/KEY: unsure
   LOCATION: (2266)...(2268)
US-09-368-590-1

Query Match         3.3%;   Score 42.8;  DB 3;  Length 7812;
Best Local Similarity  47.9%;  Pred. No. 0.0063;
Matches 156;  Conservative  0;  Mismatches 167;  Indels  3;  Gaps  1;

>y   301  CCAGGAACTTCAAGAGGCCACCGATGAGCTGACCTCCAAGCTGCGCCAAGCTGAGGTGAT  360
            ||||| | ||||  || ||||  |||| | || |||| |||||
>b  3639  CAAGGAGTTGCACCAGGTTGGCGCACGACCTGGACGACGAGGCTGGCATGGGTTCAGGAGCG  3698

>y   361  CAAGGGA---TCCTGGCAGCCCGTGGGCGATCTCCCATTGACTCTCTCCAAGATCACCT  417
            |||| |     |||| | | |||||  |||| | ||||||
>b  3699  GCTGCCACTGGCCATGCAGAGCAGACAGAGAGCGGTTCCAGCGCCGTCCAGCAGCACAT  3758

>y   418  CGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTCGAAAGAGAACGTGAGCCACGT  477
            .| || | || ||  ||  |||| || |  | |||
>b  3759  CAAAAAGAACCAGGGCCTGCGCGGGGGAGATCCAGGCGCCATGGGCCCGCCTTCACCGTAACCTCAG  3818

>y   478  CAATGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAG  537
            ||| ||||||||| | |||  | |  |||||
>b  3819  GCTGGAGCGGCGCGGGGCGCGCTGGCGTTGCGTGGGAGGCTGGCGCAGAGGCAGTGCGCCG  3878

>y   538  CACTCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCTCGAGGACCGAGT  597
            ||||||||| | | |||| |||||||
>b  3879  GGGCCTGGAGCAGCTGCGAGAGCGGCCCTGGCCCGGACTGCCGGGGAGGGCTGCCGACGGCA  3938

>y   598  CAGGCAGCTGCATGAAGCCCACAGGG  623
            ||| |||| || ||| ||
>b  3939  GCAGGTGCTGGACGCGCCGCCTTCCAGG  3964
```

```
RESULT 12
US-08-425-069-3
; Sequence 3, Application US/08425069
; Patent No. 5728810
   GENERAL INFORMATION:
   APPLICANT: Lewis, Randolph V.
   APPLICANT: Xu, Ming
   APPLICANT: Hinman, Michael B.
   TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
   TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
   TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
   NUMBER OF SEQUENCES: 69
   CORRESPONDENCE ADDRESS:
   ADDRESSEE: Birch, Stewart, Kolasch & Birch
   STREET: 301 No. 572881Oth Washington Street
   CITY: Falls Church
   STATE: Virginia
   COUNTRY: U.S.A.
   ZIP: 22046
   COMPUTER READABLE FORM:
   MEDIUM TYPE: Floppy disk
   COMPUTER: IBM PC compatible
   OPERATING SYSTEM: PC-DOS/MS-DOS
   SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/08/425,069
   FILING DATE: 19-APR-1995
   CLASSIFICATION: 435
   ATTORNEY/AGENT INFORMATION:
   NAME: Murphy Jr., Gerald M
   REGISTRATION NUMBER: 28,977
   REFERENCE/DOCKET NUMBER: 1447-106P
   TELECOMMUNICATION INFORMATION:
   TELEPHONE: (703) 205-8000
   TELEFAX: (703) 205-8050
   TELEX:
INFORMATION FOR SEQ ID NO: 3:
   SEQUENCE CHARACTERISTICS:
   LENGTH: 1995 base pairs
   TYPE: nucleic acid
   STRANDEDNESS: single
   TOPOLOGY: linear
   MOLECULE TYPE: cDNA
   HYPOTHETICAL: NO
   IMMEDIATE SOURCE:
   CLONE: p6B
   FEATURE:
   NAME/KEY: CDS
   LOCATION: 1..1785
US-08-425-069-3

Query Match         3.0%;   Score 38.6;  DB 1;  Length 1995;
Best Local Similarity  44.0%;  Pred. No. 0.061;
Matches 164;  Conservative  0;  Mismatches 209;  Indels  0;  Gaps  0;

Qy    11  GCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGGAGGAATTGAAAACTAAAGAA  70
            || | | ||||||  |  |||  ||| | ||||||
Db   731  GCAGTGCAGCTGCAGCGCCGAGCAGCAGCCGGAGGATATGGAC  790

Qy    71  CCTGTAATCATGAGTACTTCTTGAGACTGTACGAAATATTTCTGACAGAGCAGCCTTTGGAA  130
            |||| |||| |||  | |  |||| || || || |
Db   791  CAGGACAACACAAGGACCATTCTGGACCCGGTAGTGCCCGCTAGCAGCCGCCGCGAGCAG  850

Qy   131  GGACTAGAGAAACTCACCAGGGAGGCCAGAGAGCTGCCTCCTCGAGGAGAGCTGGAGAAT  190
            || | |||  || |||||||
Db   851  GACCTGGAGGATATGGCCCTGGACAACAAGGACACCGGACCTGGACAACAAG  910

Qy   191  GTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAAAATTGAAC  250
            |||| |  | |||| |||||  ||||  | |||
Db   911  GACCATCTGGAGCAGGCAGTGCAGCAGCCGCAGCAGCCCTGGACAGAACCAGGAT  970

Qy   251  CTGCACTCCGCTGACTGCGCAGAGAAAATAGATGAGACCCTGAAAGACTTCAGGAACTT  310
            ||| || | ||| |||  | |  || |||  |||
```

Db   971 TAGGAGGTTATGGACCAGGACAACAAGGTCCAGGAGGACAACAAGGTC 1030

Qy   311 CAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCC 370
               |||||| |||| ||||||| |||||||| ||||||||
Db   1031 CAGGAGGATATGGACCAGTAGTGCATCTGCAGCAGCGCCAGCAGGACCTGGACAAC 1090

Qy   371 TGGCAGCCCGTGG 383
               ||||||
Db   1091 AAGGACCAGGAGG 1103

RESULT 13
US-08-317-844B-3
; Sequence 3, Application US/08317844B
; Patent No. 5989894
; GENERAL INFORMATION:
;   APPLICANT: Lewis, Randolph V.
;   APPLICANT: Xu, Ming
;   APPLICANT: Hinman, Michael B.
;   TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
;   TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
;   TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
;   NUMBER OF SEQUENCES: 62
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Birch, Stewart, Kolasch & Birch
;   STREET: 301 No. 5989894th Washington Street
;   CITY: Falls Church
;   STATE: Virginia
;   COUNTRY: U.S.A.
;   ZIP: 22046
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/317,844B
;   FILING DATE: 04-OCT-1994
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Murphy Jr., Gerald M
;   REGISTRATION NUMBER: 28,977
;   REFERENCE/DOCKET NUMBER: 1447-105P
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703) 241-1300
;   TELEFAX: (703) 241-2848
;   TELEX: 248345
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1995 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   IMMEDIATE SOURCE:
;   CLONE: p6B
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1785
US-08-317-844B-3

Query Match   3.0%; Score 38.6; DB 2; Length 1995;
Best Local Similarity   44.0%; Pred. No. 0.061;
Matches 164; Conservative   0; Mismatches 209; Indels   0; Gaps   0;

Qy   11 GCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAA 70
               ||||| ||| || ||||| ||||||||| ||||||||
Db   731 GCAGTTCGCAGCTGCAGCCGCGACCGCCGAGCAGGACCAACAAGGACCCGGAGGATATGGAC 790

Qy   71 CCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAA 130
               |||||| || || |||||||||||
Db   791 CAGGAACAACAAGGACCATCTGGACCCGGTAGTGCCGCGCTGACACTCTGGACCGCCGCAGCAG 850

RESULT 15
US-09-668-313A-3
; Sequence 3, Application US/09668313A

Db   131 GGACTAGAGAAACTTCTACCAGGAGCCCAGAGAGAGCCTCGCTCCTCGGAGGAGAGAGCCCAGAAT 190

Db   851 GACCTGGGAGGATATGGCCCTGGACAACAAGGACCCGGAGGATATGGACCAGGACAACAAG 910

Qy   191 GTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAAC 250
               ||||||
Db   911 GACCATCTGGAGCCGGTGGCAGTGCGCAGGACGCCGAGCAGCAGCCGCGACCAACAAGGAT 970

Qy   251 CTGACTCCGCTGACTGGCAGCAGAAAATAGATGAGACCCTGAAAGACTCCAGGAACTT 310
               |||||||
Db   971 TAGGAGGTTATGGACCAGGACAACAAGGTCCAGGAGGATATGGACCAGGACAACAAGGTC 1030

Qy   311 CAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCC 370
               |||||||
Db   1031 CAGGAGGATATGGACCAGTAGTGCATCTGCAGCAGCGCCAGCAGGACCTGGACAAC 1090

Qy   371 TGGCAGCCCGTGG 383
               ||||||
Db   1091 AAGGACCAGGAGG 1103

RESULT 14
US-09-220-132-24
; Sequence 24, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
;   APPLICANT: Shyjan, Andrew W.
;   TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
;   TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
;   FILE REFERENCE: 07334-074001
;   CURRENT APPLICATION NUMBER: US/09/220,132
;   CURRENT FILING DATE: 1998-12-23
;   PRIOR APPLICATION NUMBER: US 60/079,303
;   PRIOR FILING DATE: 1998-03-25
;   PRIOR APPLICATION NUMBER: US 60/068,821
;   PRIOR FILING DATE: 1997-12-24
;   NUMBER OF SEQ ID NOS: 191
;   SEQ ID NO 24
;   LENGTH: 7672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-220-132-24

Query Match   3.0%; Score 38.4; DB 4; Length 7672;
Best Local Similarity   47.2%; Pred. No. 0.17;
Matches 117; Conservative   0; Mismatches 131; Indels   0; Gaps   0;

Qy   375 AGCCGGTGGCGATCTCCATTGACTCTCTCCAAGATCACCTCGAGAAAGTCAAGGCAC 434
               ||||||
Db   6431 ACCCCCGTGCGTGCAACTCCTTGGAGAAATCAAAGCTTTGCGCGAGGCCCACCACGCCT 6490

Qy   435 TTCGGAGGAGAAATTGCGCCTCTGAAAGAAGAACGTGAGCCACGTCAATGACCTTGCTCGCC 494
               |||||
Db   6491 TCCGGTCTCCCTCCCAGCTCTTGCCCAGGCTGACTTCAACCAGCTGGCCAGCTCTGGAAGACCTGA 6550

Qy   495 AGCTTACCACTTTGGGCATTCAGCTTCACCGTATAACCTCAGCACTCTGGAAGACCTGA 554
               ||||||
Db   6551 AGATCAAGAGCTTCCGGTAGCCTTCCAACCCTACCACCTTCGGTTTACCATGGAGGCCCCCTGG 6610

Qy   555 ACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTCCATGAAG 614
               ||||||
Db   6611 AGGAGACCTGGAGAACCTACAGGAACTACAGAGAAAATCATCAAGGAGGAGGAGCTGCAGAAGG 6670

Qy   615 CCCACAGG 622
               ||||||
Db   6671 AACAGCGG 6678

```
  Patent No. 6503756
; GENERAL INFORMATION:
;  APPLICANT:  Brett P. Monia
;  APPLICANT:  Susan M. Freier
;  APPLICANT:  Jacqueline Wyatt
;  TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
;  FILE REFERENCE: RTS-0127
;  CURRENT APPLICATION NUMBER: US/09/668,313A
;  CURRENT FILING DATE:  2000-09-22
;  NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 3
;  LENGTH: 428
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)...(171)
US-09-668-313A-3
```

```
Query Match      2.9%;    Score 38.2;  DB 4;  Length 428;
Best Local Similarity   55.7%;   Pred. No. 0.03;
Matches   73;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy    674 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACA 733
          ||||| || |||  | |   |   ||   |||| || ||  | |||| || | ||  ||
Db     13 TGGGAGGAAGCTTACACAGCAGATGGAATCAAGTACTTCATCAATCATGTAACACAGACT 72

Qy    734 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAAT 793
          |  | ||||||| |||||||   |  | ||| ||   | ||   || |||||| ||||
Db     73 ACATCCTGGATCCATCCCGTGATGAGTGTCCTGAATCTATCTCGCTCAGAGGAGAATGAA 132

Qy    794 GTCAGATTCTC 804
          | | ||||  |
Db    133 GAGGATTGCTC 143
```

Search completed: April 5, 2004, 08:24:56
Job time : 65.0939 secs

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004  Compugen Ltd.

DM nucleic - nucleic search, using sw model

un on:     April 5, 2004, 04:53:30 ; Search time 388.214 Seconds
                                     (without alignments)
                                     15454.047 Million cell updates/sec

itle:      US-09-845-416-14_COPY_900_2500
erfect score: 1601
equence:   1 gagctatgcctacacacagg........taccagtcttagctgacct 1601

coring table: IDENTITY.NUC
           Gapop 10.0 , Gapext 1.0

earched:   2470430 seqs, 1873665578 residues

otal number of hits satisfying chosen parameters:  4940860

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
           Listing first 45 summaries

Published Applications NA:*
 1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

esult No.	Score	% Query Match	Length	DB	ID	Description
1	1601	100.0	3446	10	US-09-845-416-14	Sequence 14, Appl
2	1590	99.3	4414	10	US-09-845-416-32	Sequence 32, Appl
3	1420.8	88.7	5417	15	US-10-149-736-39	Sequence 39, Appl
4	1193	74.5	3510	10	US-09-845-416-12	Sequence 12, Appl
5	1193	74.5	4476	10	US-09-845-416-31	Sequence 31, Appl
6	1102	68.8	4182	10	US-09-845-416-2	Sequence 2, Appli
7	1102	68.8	5149	10	US-09-845-416-27	Sequence 27, Appl
8	1091	68.1	8689	15	US-10-149-736-42	Sequence 42, Appl
9	1091	68.1	11443	15	US-10-149-736-44	Sequence 44, Appl
10	1091	68.1	12057	15	US-10-149-736-47	Sequence 47, Appl
11	1091	68.1	13957	9	US-09-782-378A-22	Sequence 22, Appl
12	1091	68.1	13957	9	US-09-880-107-2284	Sequence 2284, Ap
13	1091	68.1	13957	15	US-09-845-416-1	Sequence 1, Appli
14	1091	68.1	14069	12	US-10-342-887-434	Sequence 434, App
15	1091	68.1	14082	12	US-10-342-887-981	Sequence 981, App
16	1091	68.1	14082	15	US-10-341-434-108	Sequence 108, App
17	1087.8	67.9	11058	10	US-09-845-416-1	Sequence 1, Appli
18	1085.8	67.8	1991	10	US-09-845-416-9	Sequence 9, Appli
19	863	53.9	3858	10	US-09-845-416-9	Sequence 9, Appli
20	863	53.9	4825	10	US-09-845-416-29	Sequence 29, Appl
21	863	53.9	4848	10	US-09-845-416-35	Sequence 35, Appl
22	863	53.9	5060	10	US-09-845-416-36	Sequence 36, Appl
23	854.2	53.4	13815	15	US-10-149-736-2	Sequence 2, Appli
24	854	53.3	5339	15	US-10-149-736-40	Sequence 40, Appl
25	851.2	53.2	5531	10	US-09-845-416-10	Sequence 10, Appl
26	851.2	53.2	4498	10	US-09-845-416-30	Sequence 30, Appl
27	766.4	47.9	3999	10	US-09-845-416-6	Sequence 6, Appli
28	766.4	47.9	4966	10	US-09-845-416-28	Sequence 28, Appl
29	766.4	47.9	4990	10	US-09-845-416-34	Sequence 34, Appl
30	765	47.8	1667	10	US-09-845-416-7	Sequence 7, Appli
31	732.2	45.7	5462	10	US-10-149-736-41	Sequence 41, Appl
32	502.6	31.4	1821	10	US-09-845-416-13	Sequence 13, Appl
33	502.6	31.4	2169	10	US-09-845-416-4	Sequence 4, Appli
34	501	31.3	1434	10	US-09-845-416-15	Sequence 15, Appl
35	450	28.1	1340	10	US-09-845-416-11	Sequence 11, Appl
36	349.4	21.8	10302	9	US-09-782-378A-23	Sequence 23, Appl
37	349.4	21.8	10302	15	US-10-149-736-3	Sequence 3, Appli
38	349.4	21.8	16531	14	US-10-101-510-667	Sequence 667, App
39	327	20.4	327	15	US-10-149-736-8	Sequence 8, Appli
40	326.2	20.4	11096	15	US-10-149-736-4	Sequence 4, Appli
41	324	20.2	324	15	US-10-149-736-33	Sequence 33, Appl
42	322	20.1	333	15	US-10-149-736-10	Sequence 10, Appl
43	321	20.0	333	15	US-10-149-736-9	Sequence 9, Appli
44	245.4	15.3	5106	12	US-10-220-120-157	Sequence 157, App
45	190.2	11.9	256	9	US-09-864-761-21956	Sequence 21956, A

ALIGNMENTS

```
RESULT 1
US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 3446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-14

Query Match        100.0%;  Score 1601;  DB 10;  Length 3446;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1601;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT   959

Qy     61  TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG    120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    960  TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTGGCAGTTCATTGATGGAGAG   1019

Qy    121  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020  TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  1079

Qy    181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
```

RESULT 2
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
;   LENGTH: 4414
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-845-416-32

Query Match   99.3%; Score 1590; DB 10; Length 4414;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1601; Conservative 0; Mismatches 0; Indels 0; Gaps 1;

```
Db   3097  CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTCGTCTCCTGGAGGTCCCTGGAGAGAGC  3156

Qy   1500  CATCTCGCCAAACAAAGTGCCTACCATATCAACCACGAGACTCAAACAACTTGCTGGGA  1559
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3157  CATCTCGCCAAACAAAGTGCCTACTACTATCAACCACGAGACTCAAACAACTTGCTGTGGGA  3216

Qy   1560  CCATCCCAAAATGACAGAGAGCTCTACCAGTCTTAGCTGACCT  1601
           |||||||||||||||||||||||||||||||||||||||||||
Db   3217  CCATCCCAAAATGACAGAGAGCTCTACCAGTCTTTAGCTGACCT  3258
```

RESULT 3
US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 5417
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-39

Query Match   88.7%;   Score 1420.8;   DB 15;   Length 5417;
Best Local Similarity   91.0%;   Pred. No. 0;
Matches 1599;   Conservative   2;   Mismatches   0;   Indels 157;   Gaps   2;

```
Qy      1  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGAGCCCATT  1158

Qy     61  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159  TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  1218

Qy    121  TGAAGTAAACCTGGACCGTTATCAAACAGGTTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219  TGAAGTAAACCTGGACCGTTATCAAACAGGTTTAGAAGAAGTATTATCGTGGCTTCTTTC  1278

Qy    181  TGCTGAGGACACATTGCAAGCAAGCAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279  TGCTGAGGACACATTGCAAGCAAGCAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA  1338

Qy    241  CCAGTTTCATACTCATGAGGGGTACATGGATTGACAGCCCATCAGGGCCGGGTTGG  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339  CCAGTTTCATACTCATGAGGGGTACATGGATTGACAGCCCATCAGGGCCGGGTTGG  1398

Qy    301  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA  1458

Qy    361  AACTGAAGTACAAGAGCAAGATGAATCTCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459  AACTGAAGTACAAGAGCAAGATGAATCTCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC  1518

Qy    421  TAGCATGGAAAAACAAAGCAATTTACATGAGTTTTAATGGATCCCAGAATC-GAAACT  479
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1519  TAGCATGGAAAAACAAAGCAATTTACATGAGTTTTAATGGATCCCAGAATCAGAAAACT  1578

Qy    480  GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAAGAACAAGGAAGAAATGGAGGAAGA  539
```

```
.b   2017  AACTGAAGTACAAGAGCAGCAGATGAATGAATCTCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC  2076
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.y    421  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTTCCAGAATC-GAAACT  479
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
.b   2077  TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTTCCAGAATCAGAAAACT  2136
.y    480  GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAAGAACAAGGAAGAAATGGAGGAAGA  539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2137  GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAAGAACAAGGAAGAAATGGAGGAAGA  2196
.y    540  GCCTCTTGGACCTGGATCTTGAAGACTACAACGCCAAGTACAACAACATAAGGTGCTTCA  599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2197  GCCTCTTGGACCTGGATCTTGAAGACTACAACGCCAAGTACAACAACATAAGGTGCTTCA  2256
.y    600  AGAAGATCTAGAACAACAAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTGGTGGTAGT  659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2257  AGAAGATCTAGAACAACAAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTGGTGGTAGT  2316
.y    660  TGATGAATCTAGTGGAGATCACGCAACTGCTCTTTGGAAGAACAACTTAAGGTATTGGG  719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2317  TGATGAATCTAGTGGAGATCACGCAACTGCTCTTTGGAAGAACAACTTAAGGTATTGGG  2376
.y    720  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT  779
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2377  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT  2436
.y    780  CCTTCTCAAAATGGCCAACGTCTTACTGACAGATCAGTGCCTTTTTAGTGCATGGCTTCAGA  839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2437  CCTTCTCAAAATGGCAACGTCTTACTGACAGATCAGTGCCTTTTTAGTGCATGGCTTCAGA  2496
.y    840  AAAAGAAGATGCAGTGAACAAGAGATTCACAACTGCTTTAAAGATCAAAATGAAAATGTT  899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2497  AAAAGAAGATGCAGTGAACAAGAGATTCACAACTGCTTTAAAGATCAAAATGAAAATGTT  2556
.y    900  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAACTGGGCCGTTTTAAAAGCGAAAAGAAAAAGCAATCCAT  959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2557  ATCAAGTCTTCAAAAACTGGCCGTTTAAAAACTGGGCCGTTTTAAAAGCGAAAAGAAAAAGCAATCCAT  2616
.y    960  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC  1019
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2617  GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC  2676
.y   1020  CCAGAGACGGAGCGAAGCATGGCTGGATAACTTTGCCCCGTGTTGGGATAATTAGTTCCAAAA  1079
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2677  CCAGAGACGGAGCGAAGCATGGCTGGATAACTTTGCCCCGTGTTGGGATAATTAGTTCCAAA  2736
.y   1080  ACTTGAAAAGAGTACAGCACAGACACCCTGAAAGACTTCAAGGAGGCCACGGA  1139
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2737  ACTTGAAAAGAGTACAGCACAGACACCCTTGAAAGACTTCAAGGAGGCCACGGA  2796
.y   1140  TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGCCAAGCTGATCAAGGGATCCTGCGCCCGTGGG  1199
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2797  TGAGCTGGACCTCAAGCTGCGCCAAGCTGAGCTGATCAAGGGATCCTGCGCCCGTGGG  2856
.y   1200  CGATCTCCTCATTGACTTCTCCAAGATCACCTGAAGAAGTCAAGGCACTTCAGGGAGA  1259
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2857  CGATCTCCTCATTGACTTCTCCAAGATCACCTGAAGAAGTCAAGGCACTTCAGGGAGA  2916
.y   1260  AATTGCGCGCTTCTGAAAGAGAACTGAGCCAGCCACGTGACCTTGCTCGCCAGCTTACCAC  1319
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2917  AATTGCGCGCTTCTGAAAGAGAACTGAGCCAGCCACGTGACCTTGCTCGCCAGCTTACCAC  2976
.y   1320  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGAAGACCTGAAGACACCAGATG  1379
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   2977  TTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGAAGACCTGAAGACACCAGATG  3036
.y   1380  GAAGCTTCTGCAGGTGGCCGTGGCCGTCGAGGACCGAGTCAGCGAGTCGCATGAGAGCCCACAGGGA  1439
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
.b   3037  GAAGCTTCTGCAGGTGGCCGTGGCCGTCGAGGACCGAGTCAGCGAGTCGCATGAGAGCCCACAGGGA  3096
.y   1440  CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTCTCGTCTCCTGGAGGTCCCTGGAGAGAGC  1499
```

```
Db  2659 GGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 2718
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1464 TCTTTCCACGTGTGTCCAGGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTCA 1523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2719 TCTTTCCACGTGTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTA 2778

Qy  1524 CTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTA 1583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2779 CTATATCAACCACGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTA 2838

Qy  1584 CCAGTCTTTAGCTGACCT 1601
          ||||||||||||||||||
Db  2839 CCAGTCTTTAGCTGACCT 2856
```

RESULT 4
US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match   74.5%;   Score 1193;   DB 10;   Length 3510;
Best Local Similarity 84.7%;   Pred. No. 0;
Matches 1411;   Conservative   0;   Mismatches   190;   Indels   64;   Gaps   4;

```
Qy     1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 959

Qy    61 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTTGGAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1019

Qy   121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAGAAGTATTATCGTGGCTTCTTTC 1079

Qy   181 TGCTGGGACACATTGCAAGCACAAGCAGGAGAGATTTCTAATGATGTGAAGTGTGAAAGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1080 TGCTGGGACACATTGCAAGCACAAGCAGCACAGAGAGATTTCTAATGATGTGAAGTGTGAAAGA 1139

Qy   241 CCAGTTTCATACTCATGAGGGGTACAGGGTACATGAGGGGTACATGAGGCCGGGGTTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1140 CCAGTTTCATACTCATGAGGGGTACATGAGGGGTACATGAGGCCCATCAGGCCGGGGTTGG 1199

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1259

Qy   361 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260 AACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1319

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  1320 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAAACT 1379

Qy   480 GAAAGAGTTGAATGACTGCTAACAAAACAGAAGAAGAAGAAAGAAAATGGAGGAAGA 539
```

```
Db  1579 GAAAGAGTTGAATGACTGCTAACAAAAACAGAGAAGAAGAACAAGGAAATGGAGGAAGA 1638
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   540 GCCTCTTGGACCTGATCTGAAGACTCAAACGCCAAGTACACACACATAGGTGCTTCA 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639 GCCTCTTGGACCTGATCTGAAGACTCAAACGCCAAGTACACACACATAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699 AGAAGATCTAGAACAAGAACAAGTCAGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGGAGATCGGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1759 TGATGAATCTAGTGGAGATCGGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   720 AGATGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGCTTTCAGA 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGCTTTCAGA 1938

Qy   840 AAAAGAAGATGCAGTGAACAAGATCACACAACTGGCTTAAAGATCAAAATGAAATGTT 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1939 AAAAGAAGATGCAGTGAACAAGATCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1998

Qy   900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGAATCTAGAAAAGAAAAGCAATCCAT 959
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058

Qy   960 GGGCAAAACTGTATTCACTCACAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2059 GGGCAAAACTGTATTCACTCACAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy  1020 CCAGAAGACGGAAGCAATGGCTGGAATAACTTGCCCGGTGTTGGGGATAATTAGTCCAAAA 1079
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119 CCAGAAGACGGAAGCAATGGCTGGAATAACTTGCCCGGTGTTGGGGATAATTAGTCCAAAA 2178

Qy  1080 ACTTGAAAGAGTACAGCACAGA------ 1102
          ||||||||||||||||||||||
Db  2179 ACTTGAAAGAGTACAGCACAGATTTCACAGGCTGTCACCCACTCAGCCATCACTAAC 2238

Qy  1103 ------ 1102
Db  2239 ACAGACAACTGTAATGGAAACTACGGTGACCACCACAGATCCTCAAGGGAACAGAGATCCTGGTAAA 2298

Qy  1103 ------C 1103
Db  2299 GCATGCTCAAGAGGAACTTCCACCACCACCCCCAAAAGAAGAGGGCAGATTACTGTGGA 2358

Qy  1104 CCTTGAAAGACTTCAGGAACTTCAAGAGGCCACTTCAGGCGACGATGAGTGGCCAAGGCTGGCCCA 1163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2359 TCTTGAAAGACTTCAGGAACTTCAAGAGGCCACTTCAGGCGACGATGAGTGGCCAAGCTGCCGCCA 2418

Qy  1164 AGCTGAGGTGATCAAGGGATCCGGCAGCCCGTGGGCGGATCTCCTCATTGACTCCTCCA 1223
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2419 AGCTGAGGTGATCAAGGGATCCGGCAGCCCGTGGGCGGATCTCCTCATTGACTCCTCCA 2478

Qy  1224 AGATCACCCTGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCCTCTGAAAGAGAACGT 1283
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2479 AGATCACCCTGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCCTCTGAAAGAGAACGT 2538

Qy  1284 GAGCCACGTCAATGAGTCCGCGCTGCTCGCGCTTACCACCTTTGGGCATTCAGCTCTCACCGTA 1343
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2539 GAGCCACGTCAATGAGTCCGCGCTTGCTCGCGCTTACCACCTTTGGGCATTCAGCTCTCACCGTA 2598

Qy  1344 TAACCTTGCAGCACTCTGGAAGACCTCGGAAGCTTCTGGAAGCTTCTGCGGTGGCCGTCGA 1403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2599 TAACCTTGCAGCACTCTGGAAGACCTCTGGAAGCTTCTGCGGGTGGCCGTCGA 2658

Qy  1404 GGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 1463
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 5
US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US2003017131 2A1
; GENERAL INFORMATION:
;   APPLICANT: XIAO, XIAO
;   TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;   TITLE OF INVENTION: THEREOF
;   FILE REFERENCE: DE1142
;   CURRENT APPLICATION NUMBER: US/09/845,416
;   CURRENT FILING DATE: 2001-04-30
;   PRIOR APPLICATION NUMBER: 60/200,777
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
;   LENGTH: 4476
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-845-416-31

Query Match   74.5%;
Best Local Similarity   84.7%;   Score 1193;   DB 10;   Length 4476;
Matches 1411;   Conservative 0;   Mismatches 190;   Indels 64;   Gaps 4;

Qy   1 GAGCTATGCCTACACACCAGGCTGCTTATGTCACCACCTCTGACCCTCACACGAGCCCATT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1656 GAGCTATGCCTACACACCAGGCTGCTTATGTCACCACCTCTGACCCTCACACGAGCCCATT 1715

Qy   61 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1716 TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1775

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1776 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAGTATTATCGTGGCTTCTTTC 1835

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1836 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGGTGAAAGA 1895

Qy   241 CCAGTTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGCCCGGTTGG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1896 CCAGTTTCATCTCATGAGGGTACATGATGGATTGACAGCCCATCAGGCCCGGTTGG 1955

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1956 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 2015

Qy   361 AACTGAAGTACAAGAGCAGATGAATCCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2016 AACTGAAGTACAAGAGCAGATGAATCCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2075

Qy   421 TAGCATGGAAAAACAAAGCAATTTACATGAGTTTAATGGATCTCCAGAATC-GAAACT 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   2076 TAGCATGGAAAAACAAAGCAATTTACATGAGTTTAATGGATCTCCAGAATCAGAAACT 2135

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAATGAGAGGAAGA 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2136 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAAATGAGGAGGAAGA 2195

Qy   540 GCCTCTTGGACCTGACCTGATCTTGAAGACTACAAACGCCAAGTACAACACAACATAAGGTGCTTCA 599
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2196 GCCTCTTGGACCTGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCA 2255

Qy   600 AGAAGATCTAGAACAGAACAAGTCAGGGTCAATTCTCTCACTCACATGTGGTGGTAGT 659
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4182
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-845-416-2

Query Match   68.8%;   Score 1102;   DB 10;   Length 4182;
Best Local Similarity   96.9%;   Pred. No. 3.9e-312;
Matches 1134;   Conservative   0;   Mismatches   35;   Indels   1;   Gaps   1;

Qy   1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCACACGGAGCCCATT 60
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTCACACGGAGCCCATT 959

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1019

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGCTTCTTTC 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGCTTCTTTC 1079

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGAGTGGAAGTGGTGAAAGA 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGAGTGGAAGTGGTGAAAGA 1139

Qy   241 CCAGTTCATCATCATGAGGGTACATGATGATTTGAATTGAATTTGAATTTGACAGCCCATCAGGGCGGGTTGG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 CCAGTTCATCATCATGAGGGTACATGATGATTTGAATTGAATTTGACAGCCCATCAGGGCCGGGGTTGG 1199

Qy   301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA 1259

Qy   361 AACTGAAGTACAAGAGCAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGGGATGGGATGCCTCAGGGTAGC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 AACTGAAGTACAAGAGCAGAGCAGATGAATCTCCTAAATTCAAGATGGGAATGGGATGCCTCAGGGTAGC 1319

Qy   421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
                ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   1320 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1379

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAAATGGAGAAGA 539
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 GAAAGAGTTGAATGACTGGCTAACAAAAAACAGAAGAAAATGGAGAAGA 1439

Qy   540 GCCTCTTGGTCCAGCATCTGATCTTGAAGACCTAAAACGCCAAGTACACAACATAAGGTGCTTCA 599
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 GCCTCTTGGTCCAGCATCTGATCTTGAAGACCTAAAACGCCAAGTACACAACATAAGGTGCTTCA 1499

Qy   600 AGAAGATCTAGAACAACAAGACAAGTCAGGTCAATTCTCTCACTGGTTGGTGGTAGT 659
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 AGAAGATCTAGAACAACAAGACAAGTCAGGTCAATTCTCTCACTCACTGGTTGGTGGTAGT 1559

Qy   660 TGATGAATCTAGTGGAGATCGAGATCTTGAAGAACAACTTGCTTGCTTTGGAAGAACAACAACTTAAGGTATTGGG 719
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1560 TGATGAATCTAGTGGAGATCGAGATCTTGAAGAACAACTTGCTTGCTTCTTTGGAAGAACAACTTAAGGTATTGGG 1619

Qy   720 AGATCGATGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTACAAGACAT 779

Db   2256 AGAAGATCTAGACAGAAGAACAAGTCAGGTCAATTCTCTCACTGGTTGGTGGTAGT 2315
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   660 TGATGAATCTAGTGGAGATCGAGATCTTGAAGAACAACTTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2316 TGATGAATCTAGTGGAGATCGAGATCTTGAAGAACAACTTGCTTGCTTTGGAAGAACAACTTAAGGTATTGGG 2375

Qy   720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTACAAGACAT 779
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2376 AGATCGATGGGCAAACATCATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTACAAGACAG 2435

Qy   780 CCTTCTCAAATGCAACGTCTTACTACTGAAGACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2436 TTCTGACCAGTGACAGCTTGCTGCACTTTCTCTGCAGGGAACTGCCTGTGTGGCTACAGCT 2495

Qy   840 AAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAAATGTT 899
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2496 GAAAGATGATGAATTGAATTAAGCGGCAGGCGGCACCCTATTGGAGGCGACTTTCCAGCAGTTCAGAA 2555

Qy   900 ATCAAGTCTTCAAAAACTGGCCCGTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2556 GCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAAAAACTAAAGAACCTGTAATCAT 2615

Qy   960 GGGCCAAACTGTA------TTCACTCAAACAAGATCTTCTTTCAACACTGAAG 1005
                ||||||||||| ||||||||||||||||||||||||||||||
Db   2616 GAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAGA 2675

Qy   1006 AATAAGTCAGTGACCCAGGAGACGGAAGCATGGCTGA------------- 1043
                  |||||||||||||||||||||||||||||||||
Db   2676 ACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGGCCCAGAATGTCACTCGGCT 2735

Qy   1044 ----TAACTTTGCCCGGTGTTGGGGATAATTTAGTCCAAAAAACTTGAA------ 1086
                       ||||||||||||||||||||||||||||||||||||||||
Db   2736 TCTACGAAAGCAGGCTGAGGAGGCTCAATACTGAGTGAGTGGGAAAAATTGAACCTGCACTCCGC 2795

Qy   1087 ----------AAGTACAGCACAGACCCTGAAAGACTCCAGGAACTTCAAGAGGCCAC 1136
                          |||||||||||||||||||||||||||||||||||||||||||||
Db   2796 TGACTGGCAGAGAGAAAATAGAGATGAGACCCTGAAAGACTCCAGGAACTTCAAGAGGCCAC 2855

Qy   1137 GGATGAGCTGGACCTCAAGGTCCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGT 1196
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2856 GGATGAGCTGGACCTCAAGGTCCAAGCTGAGGTGATCAAGGATCCTGCAGCCCGT 2915

Qy   1197 GGGGCCGATCCCTCATTGACTCCTCCAAGATCACCTCGAGAAAGTCAGGGCACTTCGAGG 1256
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2916 GGGGCCGATCCCTCATTGACTCCTCCAAGATCACCTCGAGAAAGTCAGGCACTTCGAGG 2975

Qy   1257 AGAAAATTGCGCCCTCGAAAGAGACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 1316
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2976 AGAAAATTGCGCCCTCGAAAGAGACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTAC 3035

Qy   1317 CACTTTGGGCATTCAGCTCAGCCCTGTCGCGGTATAACCTTCAGCACTCTGGAAGACTTGAACACCAG 1376
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3036 CACTTTGGGCATTCAGCTCAGCCCTGTCGCGGTATAACCTTCAGCACTCTGGAAGACTTGAACACCAG 3095

Qy   1377 ATGGAAGCTTCTGCGAGGTGCCGTCGAGGACGAGTCAGCAGCTGCATGAGCCCCACAG 1436
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3096 ATGGAAGCTTCTGCGAGGTGCCGTCGAGGACGAGTCAGCAGCTGCATGAGCCCCACAG 3155

Qy   1437 GGACTTTGGTCCAGCATCTTCTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 1496
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3156 GGACTTTGGTCCAGCATCTTCTTCTTCCACGTCTGTCCAGGGTCCCTGGGAGAG 3215

Qy   1497 AGCCATCTCGCCAAAAAGTGCCCTACTATATCAACCAGAGACTCAAACAACTTGCTG 1556
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3216 AGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACCACGAGACTCAAACAACTTGCTG 3275

Qy   1557 GGACCATCCCAAATGACAGAGCTCTACCAGTCTTAGCTGACCT 1601
                ||||||||||||||||||||||||||||||||||||||||||
Db   3276 GGACCATCCCAAATGACAGAGCTCTACCAGTCTTAGCTGACCT 3320

RESULT 6

```
Qy    241  CCAGTTTCATACTCATGAGGGTACACTGATGGATTTGACAGCCCATCAGGCCGGGTTGG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897  CCAGTTTCATACTCATGAGGGGTACACTGATGGATTTGACAGCCCATCAGGCCGGGTTGG 1956

Qy    301  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATATCAGAAGATGAAGA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957  TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAAATATCAGAAGATGAAGA 2016

Qy    361  AACTGAAGTACAAGAGCAGAGCAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017  AACTGAAGTACAAGAGCAGAGCAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 2076

Qy    421  TAGCATGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATC-GAAACT  479
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2077  TAGCATGAAAAACAAAGCAATTACATAGAGTTTAATGATCTCCAGAATCAGAAAACT 2136

Qy    480  GAAAGAGTGAATGACTGGCTAACAAAAACAGAAGAACACGAAAATGGACGGAGA  539
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137  GAAAGAGTGAATGACTGGCTAACAAAAAACAGAAGAACACAAGGAAAATGGAGGAAGA 2196

Qy    540  GCCTCTTGGACCTGATCTTGAAGAACTTAAAACGCCAAGTACAACAACATAAGGTGCTTCA  599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197  GCCTCTTGGACCTGATCTTGAAGAACTTAAAACGCCAAGTACAACAACATAAGGTGCTTCA 2256

Qy    600  AGAAGATCTAGAACAAGAAGAACAAGTCAGGTCAATTCTCACTCACATGGTGGTGGTAGT  659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257  AGAAGATCTAGAACAAGAAGAACAAGTCAGGGTCAATTCTCTCACTCACATGGTGGTGGTAGT 2316

Qy    660  TGATGAATCTAGTGGAGATCACGCAACTGCTCTTGGAAGAACACCGCTGGGTTCTTTTACAAGACAT  719
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317  TGATGAATCTAGTGGAGATCACGCAACTGCTCTTGGAAGAACACCGCTGGGTTCTTTTACAAGACAT 2376

Qy    720  AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGGCTGGGTTCTTTTTACAAGACAT  779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377  AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGGCTGGGTTCTTTTTACAAGACAT 2436

Qy    780  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA  839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2437  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 2496

Qy    840  AAAAGAAGATGCAGTGAACACAGATTCACACAACTCGGCTTTAAAGATCAAAATGAAATGTT  899
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2497  AAAAGAAGATGCAGTGAACACAGATTCACACAACTCGGCTTTAAAGATCAAAATGAAATGTT 2556

Qy    900  ATCAAGTCTTCAAAAACTGGCCGTTTCAAACAAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC  959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2557  ATCAAGTCTTCAAAAACTGGCCGTTTCAAACAAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC 2616

Qy    960  GGGCAAACTGTATTCACTCAAACAAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2617  GGGCAAACTGTATTCACTCAAACAAGATCTCTTCAACACTGAAGAATAAGTCAGTGAC 2676

Qy   1020  CCAGAAGACGGAAGCATGGCATGGCCTGGATAACTTTGCCCGGTGTGTTGGGATAATTTAGTCCAAAA 1079
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2677  CCAGAAGACGGAAGCATGGCATGGCCTGGATAACTTTGCCCGGTGTGTTGGGATAATTTAGTCCAAAA 2736

Qy   1080  ACTTGAAAAGAGTACAGCACAGACCCTGAAAGACTCCAGGAACTTCAGAGAGGCCACGGA 1139
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2737  ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCGTGGACCTGGA 2796

Qy   1140  TGAGCTGGACCTCAAGCTGCGCCAAGCTGA 1169
           ||||||||||||||||||||||||||||||
Db   2797  AAAGTTTCTTGCTGCTTACAGAAGCTGA 2826
```

RESULT 8

```
Db   1620  AGATCGATGGGCAAAACATCTGTAGATGGACAGAAGACCGGCTGGGTTCTTTTTACAAGACAT 1679

Qy    780  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA  839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1680  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 1739

Qy    840  AAAAGAAGATGCAGTGAACACAGATTCACACAACTCGGCTTTAAAGATCAAAATGAAATGTT  899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1740  AAAAGAAGATGCAGTGAACACAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1799

Qy    900  ATCAAGTCTTCAAAAACTGGCCGTTTCAAACAAGATCTCTTCAACACTGAAGAATAAGCAATCCAT  959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1800  ATCAAGTCTTCAAAAACTGGCCGTTTCAAACAAGATCTCTTCAACACTGAGAAAAGAAAAGCAATCCAT 1859

Qy    960  GGGCAAACTGTATTCACTCAAACAAGATCTCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1860  GGGCAAACTGTATTCACTCAAACAAGATCTCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1919

Qy   1020  CCAGAAGACGGAAGCATGGCTGATGAATAACTTTGCCCCGTGTTGGGGATAAATTTAGTCCAAAA 1079
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1920  CCAGAAGACGGAAGCATGGCTGATGAATAACTTTGCCCCGTGTTGGGGATAAATTTAGTCCAAAA 1979

Qy   1080  ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 1139
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1980  ACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAACAGTTCCCCCTGGACCTGGA 2039

Qy   1140  TGAGCTGGACCTCAAGCTGCGCCAAGCTGA 1169
           ||||||||||||||||||||||||||||||
Db   2040  AAAGTTTCTTGCTGCTTACAGAAGCTGA 2069
```

```
Query Match    68.8%;  Score 1102;  DB 10;  Length 5149;
Best Local Similarity   96.9%;  Pred. No. 4.5e-312;
Matches 1134;  Conservative    0;  Mismatches   35;  Indels    1;  Gaps    1;

Qy      1  GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657  GAGCTATGCCTACACACAGGCTGCTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1716

Qy     61  TCCTTCACAGCATTTGGAAGCTCCGAAGCAAGTCATTTGGCAGTTCATTGATGGAGAG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717  TCCTTCACAGCATTTGGAAGCTCCGAAGCAAGTCATTTGGCAGTTCATTGATGGAGAG 1776

Qy    121  TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777  TGAAGTAAACCTGGACCCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1836

Qy    181  TGCTGAGGACACATTCGAAGCACAAGCACAAGGAGACGAATTTCTAATGATGTGGAAGTGGTGAAAGA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837  TGCTGAGGACACATTCGAAGCACAAGCACAAGGAGACGAATTTCTAATGATGTGGAAGTGGTGAAAGA 1896
```

```
 ;   FILE REFERENCE: UM-06968
 ;   CURRENT APPLICATION NUMBER: US/10/149,736
 ;   CURRENT FILING DATE: 2002-06-17
 ;   PRIOR APPLICATION NUMBER: PCT/US01/31126
 ;   PRIOR FILING DATE: 2001-10-04
 ;   PRIOR APPLICATION NUMBER: 60/238,848
 ;   PRIOR FILING DATE: 2000-10-06
 ;   NUMBER OF SEQ ID NOS: 96
 ;   SOFTWARE: PatentIn version 3.1
 ;   SEQ ID NO 42
 ;   LENGTH: 8689
 ;   TYPE: DNA
 ;   ORGANISM: Artificial Sequence
 ;   FEATURE:
 ;   OTHER INFORMATION: Synthetic
US-10-149-736-42

Query Match         68.1%;   Score 1091;   DB 15;   Length 8689;
Best Local Similarity   99.9%;   Pred. No. 1.1e-308;
Matches 1102;   Conservative  0;   Mismatches  0;   Indels  1;   Gaps  1;

Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT 1158

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTC 1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 1338

Qy    241 CCAGTTTCATACTTCATGAGGGGTACAGCTGATTTGACAGCCCATCAGGGGCCGGGTTGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTTCATGAGGGGTACAGCTGATTTGACAGCCCATCAGGGGCCGGGTTGG 1398

Qy    301 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy    361 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy    421 TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1519 TAGCATGGAAAAACAAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy    480 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGAGGAAGA 539
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAGAACAAGGAAAATGAGGAAGA 1638

Qy    540 GCCTCTTGGACCTGATCTTGAAGACTTGAAGACTTGAAGACTACAACAACATAAGGTGCTTCA 599
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACTTGAAGACTTGAAGACTACAACAACATAAGGTGCTTCA 1698

Qy    600 AGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATCTCTCACTCACATGGTGGTAGT 659
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGAACAAGTCAGGGTCAATCTCTCACTCACATGGTGGTAGT 1758

Qy    660 TGATGAATCTCAGTGGAGATCACGCAACTCTGCTGCTTTGGAAGACAACTCAAGGTATTGGG 719
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTCAGTGGAGATCACGCAACTCTGCTGCTTTGGAAGACAACTCAAGGTATTGGG 1818

Qy    720 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878
```

```
Qy    780 CCTTCTCAAATGCAACGTCTTACTGAAGAACACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy    840 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 899
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 1998

Qy    900 ATCAAGTCTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 959
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTCAAAAACTGGCCGTTTAAAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058

Qy    960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy   1020 CCAGAGACGGAAGCATGGCTGGATAACTTGCGTTGGGATATAATTAGTCCAAAA 1079
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAA 2178

Qy   1080 ACTTGAAAAGAGTACAGCACAGA 1102
            |||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGA 2201
```

```
RESULT 9
US-10-149-736-44
; Sequence 44, Application US/10149736
; Publication No. US2003021633A1
;   GENERAL INFORMATION:
 ;   APPLICANT: Chamberlain, Jeffrey S.
 ;   APPLICANT: Harper, Scott Q.
 ;   TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
 ;   FILE REFERENCE: UM-06968
 ;   CURRENT APPLICATION NUMBER: US/10/149,736
 ;   CURRENT FILING DATE: 2002-06-17
 ;   PRIOR APPLICATION NUMBER: PCT/US01/31126
 ;   PRIOR FILING DATE: 2001-10-04
 ;   PRIOR APPLICATION NUMBER: 60/238,848
 ;   PRIOR FILING DATE: 2000-10-06
 ;   NUMBER OF SEQ ID NOS: 96
 ;   SOFTWARE: PatentIn version 3.1
 ;   SEQ ID NO 44
 ;   LENGTH: 11443
 ;   TYPE: DNA
 ;   ORGANISM: Artificial Sequence
 ;   FEATURE:
 ;   OTHER INFORMATION: Synthetic
US-10-149-736-44

Query Match         68.1%;   Score 1091;   DB 15;   Length 11443;
Best Local Similarity   99.9%;   Pred. No. 1.3e-308;
Matches 1102;   Conservative  0;   Mismatches  0;   Indels  1;   Gaps  1;

Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCACCACTTCTGACCCTACACGGAGCCCATT 1158

Qy     61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTCATTGATGAGGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTCATTGATGAGGAG 1218

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy    181 TGCTGAGGACACATTGCAAGCTCCTGAAGCTCCTGAAGCTCATTTGGCAGTCATTTGGCAGTCATTTGGCAGTGTGAAAGA 1218
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TCCTTCACAGCATTTGGAAGCTCCTGAAGCTCCTGAAGCTCATTTGGCAGTGTGAAAGA 1218

Qy    121 TGAAGTAAACTGGACCGTTATCAAACAGCACAAGCACAAGCTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACTGGACCGTTATCAAACAGCACAAGCACAAGCTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGCTTCATTTGGCAGTCATTTGGCAGTGATGGAGAG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGCTTCATTTGGCAGTCATTTGGCAGTGATGGAGAG 1338

Qy    181 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTCATTGATGGAGAG 1218

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGCACAAGCACAAGCTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGACCGTTATCAAACAGCACAAGCACAAGCTTAGAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGCAAGCTTCATTTGGCAGTCATTTGGCAGTGATGGAGAG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGCAAGCTTCATTTGGCAGTCATTTGGCAGTGATGGAGAG 1338

Qy    241 CCAGTTTCATACTTCATGAGGGGTACATGATGATTTGCAGCGTACATGATGATTTGCAGCCCATCAGGCCGGGTTGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTTCATGAGGGGTACATGATGATTTGCAGCCCATCAGGCCGGGTTGG 1398
```

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 12057
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-47

Query Match   68.1%;   Score 1091;   DB 15;   Length 12057;
Best Local Similarity 99.9%;   Pred. No. 1.3e-308;
Matches 1102;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   1;

```
Qy     1 GAGCTATGCCTACACACGGCTGCTATGCCACCTCTGACCACCTCTGACCCTACACGGAGCCCATT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 GAGCTATGCCTACACACGGCTGCTATGTCACCACCTCTGACCCCTACACGGAGCCCATT 1158

Qy    61 TCCTTCACAGATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

Qy   121 TGAAGTAAACCTGGACCGTTATCAAACAGTTTAGAGAAGAGTATTATCGTGGCTTCTTTC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 TGAAGTAAACCTGGACCGTTATCAAACAGTTTAGAGAAGAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGTGGAAGTGGTGAAAGA    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTCATCTCATGAGGGTACATGATGGTTACAGCCCATCAGGGCCGGGTTGG       300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1339 CCAGTTCATCTCATGAGGGGTACATGATGGGTTACAGCCCATCAGGGCCGGGTTGG     1398

Qy   301 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACCAGGAAAATTATCAGAAGATGAAGA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1399 TAATATTCACAATTGGGAAGTAAGCTGATTGGAACCAGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAGATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1459 AACTGAAGTACAAGAGCAGATGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC   1518

Qy   421 TAGCATGGAAAAACAAAGCAATTACACAGAGTTTAATGAATCTCCAGAATC-GAAACT    479
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  1519 TAGCATGGAAAAACAAAGCAATTTACACAGAGTTTAATGAATCTCCAGAATCAGAAACT 1578

Qy   480 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA   539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579 GAAAGAGTTGAATGACTGGCTAACAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA  1638

Qy   540 GCCTCTTGGACCTGATCTTGAAGACTTAAGGATTGCTTCA                     599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1639 GCCTCTTGGACCTGATCTTGAAGACTTAAGGATTGCTTCA                    1698

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGTGGTGGTAGT 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCCACTCACATGGTGTGGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG    719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1759 TGATGAATCTAGTGAGATCACGCAACTGCTGCCTTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   720 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTTACAAGACAT 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTTACAAGACAT 1878

Qy   780 CCTTCTCAAATGGCAACGTCTTACTCAGAAGACAGTGCCTTTTAGTGCATGGCTTTCAGA  839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1879 CCTTCTCAAATGGCAACGTCTTACTCAGAAGACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   840 AAAAGAAGATGCAGTGACAAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT    899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1939 AAAAGAAGATGCAGTGACAAGATTCACACAACTGGCTTTAAAGATCAAAATGAAATGTT  1998

Qy   900 ATCAAGTCTTCAAAACTGGGCCGTTTAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT  959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1999 ATCAAGTCTTCAAAACTGGGCCGTTTAAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058

Qy   960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy  1020 CCAGAGACGGAAGCATGGCCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAA   1079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119 CCAGAGACGGAAGCATGGCCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAA   2178

Qy  1080 ACTTGAAAAGAGTACAGCACAGA 1102
          |||||||||||||||||||||||
Db  2179 ACTTGAAAAGAGTACAGCACAGA 2201
```

Qy   421 TAGCATGGAAAACAAAGCAATTACATAGAGTTTTAATGGATCTCCAGAATC-GAAACT 479
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   1519 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy   480 GAAAGAGTTGAATGACTGGCTACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA 539
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGA 1638

Qy   540 GCCTCTTGACCTGATCTTGAAGACCTTAAAAGCCAAGTACAACACATAAGGTGCTTCA 599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGACCTGATCTTGAAGACCTTGAAGACCTTAAAAGCCAAGTACAACACATAAGGTGCTTCA 1698

Qy   600 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTTAAGGTGGTGTAGT 659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCTCACTCACTCACATGGTGGTGTGGTAGT 1758

Qy   660 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 719
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy   720 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 1878

Qy   780 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy   840 AAAAGAGATGCAGTGAACAAGATTCACCAACTGGCTTAAAGATCAAAATGAAATGTT 899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAAGATTCACCAACAACTGGCTTAAAGATCAAAATGAAATGTT 1998

Qy   900 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAACGGATCTAGAAAAAGAAATCCAT 959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAACGGATCTAGAAAAAGAAAAAGCAATCCAT 2058

Qy   960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACAAGATCTTCTTCAACAGTCAGTGAC 1019
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 2118

Qy   1020 CCAGAGACGGAAGCAGTACAGCACACAGA 1102
           |||||||||||||||||||||||||||||
Db   2119 CCAGAGACGGAAGCAGTACAGCACACAGA 2201

Qy   1080 ACTTGAAAAGAGTACAGCACACAGA 1102
           |||||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACACAGA 2201

RESULT 12
US-09-880-107-2284
; Sequence 2284, Application US/09880107
; Patent No. US2002014298lAl
; GENERAL INFORMATION:
;   APPLICANT: Horne, Darci T.
;   APPLICANT: Vockley, Joseph G.
;   APPLICANT: Scherf, Uwe
;   APPLICANT: Gene Logic, Inc.
;   TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;   FILE REFERENCE: 44921-5028-WO
;   CURRENT APPLICATION NUMBER: US/09/880,107
;   CURRENT FILING DATE: 2001-06-14
;   PRIOR APPLICATION NUMBER: US 60/211,379
;   PRIOR FILING DATE: 2000-06-14
;   PRIOR APPLICATION NUMBER: US 60/237,054
;   PRIOR FILING DATE: 2000-10-02
;   NUMBER OF SEQ ID NOS: 3950
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2284
;   LENGTH: 13957
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:

Qy   900 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGGCGGATCTAGAAAAAGAAAAAGCAATCCAT 959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGGCGGATCTAGAAAAAGAAGAATAAGTCAGTGAC 2058

Qy   960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC 1019
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAATAGTCAGTGAC 2118

Qy   1020 CCAGAGACGGAAGCAGTACAGCACACAGA 1102
           |||||||||||||||||||||||||||||
Db   2179 CCAGAGACGGAAGCAGTACAGCACACAGA 2201

RESULT 11
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US2002010273lAl
; GENERAL INFORMATION:
;   APPLICANT: Hearing, Patrick
;   APPLICANT: Bahou, Wadie
;   APPLICANT: Sandalon, Ziv
;   APPLICANT: Gnatenko, Dmitri
;   TITLE OF INVENTION: Adenoviral Vectors
;   FILE REFERENCE: STONYB-04970
;   CURRENT APPLICATION NUMBER: US/09/782,378A
;   CURRENT FILING DATE: 2001-02-12
;   PRIOR APPLICATION NUMBER: 60/237,747
;   PRIOR FILING DATE: 2000-10-02
;   NUMBER OF SEQ ID NOS: 27
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 13957
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-782-378A-22

Query Match   68.1%; Score 1091; DB 9; Length 13957;
Best Local Similarity   99.9%; Pred. No. 1.5e-308;
Matches 1102; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

Qy   1 GAGCTACGCCTACACACAGGCTGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1158

Qy   61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1218

Qy   121 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGACCGGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTC 1278

Qy   181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy   241 CCAGTTTCATACTTCATGAGGGTACATGAGTTGACAGCCCATCAGGGGCCGGGTTGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTTCATGAGGGTACATGAGTTGACAGCCCATCAGGGGCCGGGTTGG 1398

Qy   301 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAATTATCAGAAGATGAAGA 1458

Qy   361 AACTGAAGTACAAGAGCAAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGCAAGATGAATCTCCTAAATTCAAGATGGGAATGCCTCAGGGGTAGC 1518

. OTHER INFORMATION: Genbank Accession No. US2002142981A1 M18533
JS-09-880-107-2284

Query Match   68.1%;   Score 1091; DB 9;   Length 13957;
Best Local Similarity 99.9%; Pred. No. 1.5e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCCGAAGACAAGTCATTCGCAGTCATTGATGGAGAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCCGAAGACAAGTCATTCGCAGTCATTGATGGAGAG 1158

Qy     61 TCCTTCACAGCATTGGAAGTCCCTGAAGACAAGTCCTTATTCGCAGTATTATCGCAGTTATCTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGTCCCTGAAGACAAGTCCTTATTCGCAGTATTATCGCAGTTATCTTTC 1218

Qy    121 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGACAAGTATTATCGTGGCTTCTTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGACAAGTATTATCGTGGCTTCTTTC 1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy    241 CCAGTTTCATACTCATGAGGGGTACATGACAGCCCATCAGGGCCGGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCATGAGGGGTACATGACAGCCCATCAGGGCCGGGTTGG 1398

Qy    301 TAATATTCTACAATTGGAAGTAAGCTGAATGGGAATGGGAATGCCTCAGGGTAGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGAAGTAAGCTGAATGGGAATGGGAATGCCTCAGGGTAGC 1458

Qy    361 AACTGAAGTACAAGAGACAGAGAATGGGAATGGGAATGCCTCAGGGTAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGACAGAGAATGGGAATGGGAATGCCTCAGGGTAGC 1518

Qy    421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy    480 ATCAAGTCTTCAAAAACTGGGGCGCGTTTAAAAGCGGATCTAGAAAGAAAAAAGCAATCCAT 539
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 ATCAAGTCTTCAAAAACTGGGGCGCGTTTAAAAGCGGATCTAGAAAGAAAAAAGCAATCCAT 1638

Qy    960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAAACTGAGAATAAGTCAGTGAC 1019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 13
US-10-149-736-1
; Sequence 1, Application US/10149736
; Publication No. US2003021633A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 13957
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-149-736-1

Query Match   68.1%;   Score 1091;   DB 15;   Length 13957;
Best Local Similarity 99.9%; Pred. No. 1.5e-308;
Matches 1102; Conservative 0; Mismatches 0; Indels 0; Gaps 1;

```
Qy      1 GAGCTATGCCTACACACAGGCTGCTTATGTCCGAAGACAAGTCATTCGCAGTCATTGATGGAGAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 GAGCTATGCCTACACACAGGCTGCTTATGTCCGAAGACAAGTCATTCGCAGTCATTGATGGAGAG 1158

Qy     61 TCCTTCACAGCATTGGAAGTCCCTGAAGACAAGTCCTTATTCGCAGTATTATCGCAGTTATCTTTC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 TCCTTCACAGCATTGGAAGTCCCTGAAGACAAGTCCTTATTCGCAGTATTATCGCAGTTATCTTTC 1218

Qy    121 TGAAGTAAACCGGTTATCAAACAGCTTAGAAGACAAGTATTATCGTGGCTTCTTTC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 TGAAGTAAACCGGTTATCAAACAGCTTAGAAGACAAGTATTATCGTGGCTTCTTTC 1278

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA 1338

Qy    241 CCAGTTTCATACTCATGAGGGGTACATGACAGCCCATCAGGGCCGGGTTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 CCAGTTTCATACTCATGAGGGGTACATGACAGCCCATCAGGGCCGGGTTGG 1398

Qy    301 TAATATTCTACAATTGGAAGTAAGCTGAATGGGAATGGGAATGCCTCAGGGTAGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 TAATATTCTACAATTGGAAGTAAGCTGAATGGGAATGGGAATGCCTCAGGGTAGC 1458

Qy    361 AACTGAAGTACAAGAGACAGAGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 AACTGAAGTACAAGAGACAGAGAATCCTAAATTCAAGATGGGAATGCCTCAGGGTAGC 1518

Qy    421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT 1578

Qy    480 GAAAGAGTTGAATGACTGGCTAACAATTGGAAGCTAAGCTAAGCTGGAAGATGAAGA 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1579 GAAAGAGTTGAATGACTGGCTAACAATTGGAAGCTAAGCTAAGCTGGAAGATGAAGA 1638

Qy    540 GCCTCTTGGACCTGACCTGAGAGAAACGGCCAAGTACAACATAAGGTGTGCTTCA 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGACCTGAGAGAAACGGCCAAGTACAACATAAGGTGTGCTTCA 1698

Qy    600 AGAAGATCTAGAACAAGAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT 1758

Qy    660 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATCTAGTGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1818

Qy    720 AGATCGATGGGCAAACATCTGTAGATCGGAGAACCGCTGGGTTCTTTTACAAGACAT 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATCGGAGAACCGCTGGGTTCTTTTACAAGACAT 1878

Qy    780 CCTTCTCAAATGGCAACGTCAACGTCTTACTGAAGACAGTGCCTTTTAGTGCATGGCTTTCAGA 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGGCAACGTCAACGTCTTACTGAAGACAGTGCCTTTTAGTGCATGGCTTTCAGA 1938

Qy    840 AAAAGAGATGCAGTGAACAAGCAGATTCACACAACTTCACACAACTGGAATGGAAATGTT 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAGATGCAGTGAACAAGCAGATTCACACAACTTCACACAACTGGAATGGAAATGTT 1998

Qy    900 ATCAAGTCTTCAAAAACTGGGGCGCGTTTAAAAGCGGATCTAGAAAGAAAAAAGCAATCCAT 959
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGGGCGCGTTTAAAAGCGGATCTAGAAAGAAAAAAGCAATCCAT 2058

Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAAACTGAGAATAAGTCAGTGAC 2118

Qy   1020 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 1079
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 2178

Qy   1080 ACTTGAAAAGAGTACAGCACAGA 1102
          |||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGA 2201
```

```
Best Local Similarity   99.9%;   Pred. No. 1.5e-308;
Matches 1102;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

Qy      1 GAGCTATGCTACACACAGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1211 GAGCTATGCTACACACAGCTGCTTATGTCACCACCTCGACCCTACACGGAGCCCATT   1270

Qy     61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1271 TCCTTCACAGCATTGGAAGCTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG   1330

Qy    121 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGCTTCTTC   180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1331 TGAAGTAAACCTGACCGTTATCAAACAGCTTTAGAGAAGAAGTATTATCGTGCTTCTTTC   1390

Qy    181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1391 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGA   1450

Qy    241 CCAGTTTCATCATCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCGCGGGTTGG   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1451 CCAGTTTCATCATCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCGCGGGTTGG   1510

Qy    301 TAATATTTCACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA   360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1511 TAATATTTCACAATTGGGAAGTAAGCTGATTGGAACAGGAGAAAATTATCAGAAGATGAAGA   1570

Qy    361 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGCGGAATGCTCCAGGGTAGC   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1571 AACTGAAGTACAAGACAGATGAATCTCCTAAATTCAAGATGCGGAATGCTCCAGGGTAGC   1630

Qy    421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATC-GAAACT   479
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   1631 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTAATGGATCTCCAGAATCAGAAACT   1690

Qy    480 GAAAGAGTTGAATGACTGCTAACAAAACAGAAGAAACAGAAAATGGAGGAAGA   539
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1691 GAAAGAGTTGAATGACTGCTAACAAAACAGAAGAAACAGAAAATGGAGGAAGA   1750

Qy    540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCA   599
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1751 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCA   1810

Qy    600 AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT   659
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1811 AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT   1870

Qy    660 TGATGAATTCAGTGGAGATCGGAGATCATCGTGAGTGGACAGAGACCGCTGGGTTCTTTTACAAGACATT   719
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1871 TGATGAATTCAGTGGAGATCGGAGATCATCGTGAGTGGACAGAGACCGCTGGGTTCTTTTACAAGACATT   1930

Qy    720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAT   779
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1931 AGATCGATGGGCAAACATCTGTAGATGGACAGAGAAGACCGCTGGGTTCTTTTACAAGACAT   1990

Qy    780 CCTTCTCAAATGCCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA   839
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1991 CCTTCTCAAATGCCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA   2050

Qy    840 AAAAGAAGATGCAGTGAACAAGATTCACCAACTGGCTTAAAGATCAAAATGAAATGTT   899
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2051 AAAAGAAGATGCAGTGAACAAGATTCACCAACTGGCTTAAAGATCAAAATGAAATGTT   2110

Qy    900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGAATCTAGAAAAAGAAAAAGCAATCCAT   959
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2111 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGAATCTAGAAAAAGAAAAAGCAATCCAT   2170

Qy    960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTCAAGAGATAAGTCAGTGAC   1019
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2171 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTCAAGAGATAAGTCAGTGAC   2230

Qy   1020 CCAGAAGACGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA   1079
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy    540 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCA   599
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCTCTTGGACCTGATCTTGAAGACCTAAAACGCCAAGTACAACACAACATAAGGTGCTTCA   1698

Qy    600 AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT   659
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 AGAAGATCTAGAACAAGACAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT   1758

Qy    660 TGATGAATTCAGTGGAGATCGGAGATCACCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG   719
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1759 TGATGAATTCAGTGGAGATCGGAGATCACCGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG   1818

Qy    720 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT   779
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1819 AGATCGATGGGCAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT   1878

Qy    780 CCTTCTCAAATGCCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA   839
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1879 CCTTCTCAAATGCCAACGTCTTACTGAAGAACAGTGCCTTTTAGTGCATGGCTTTCAGA   1938

Qy    840 AAAAGAAGATGCAGTGAACAAGATTCACCAACTGGCTTTAAAGATCAAAATGAAATGTT   899
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1939 AAAAGAAGATGCAGTGAACAAGATTCACCAACTGGCTTTAAAGATCAAAATGAAATGTT   1998

Qy    900 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGAATCTAGAGAAAAGAAAAAGCAATCCAT   959
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1999 ATCAAGTCTTCAAAAACTGGCCGTTTAAAAGCGGAATCTAGAGAAAAGAAAAAGCAATCCAT   2058

Qy    960 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC   1019
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2059 GGGCAAACTGTATTCACTCAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAGTGAC   2118

Qy   1020 CCAGAAGACCGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA   1079
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 CCAGAAGACCGGAAGCATGGCTGGATAACTTGCCCGGTGTTGGGATAATTTAGTCCAAAA   2178

Qy   1080 ACTTGAAAAGAGTACAGCACAGA   1102
            |||||||||||||||||||||||
Db   2179 ACTTGAAAAGAGTACAGCACAGA   2201
```

```
RESULT 14
US-10-342-887-434
; Sequence 434, Application US/10342887
; Publication No. US20040058340A1
GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 434
; LENGTH: 14069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-434

Query Match      68.1%;   Score 1091;   DB 12;   Length 14069;
```

```
Db     2231  CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 2290

Qy     1080  ACTTGAAAAGAGTACAGCACAGA 1102
             |||||||||||||||||||||||
Db     2291  ACTTGAAAAGAGTACAGCACAGA 2313


RESULT 15
JS-10-342-887-981
; Sequence 981, Application US/10342887
; Publication No. US2004005834OA1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 981
; LENGTH: 14082
; TYPE: DNA
; ORGANISM: Homo sapiens
JS-10-342-887-981

Query Match    68.1%;  Score 1091;  DB 12;  Length 14082;
Best Local Similarity  99.9%;  Pred. No. 1.5e-308;
Matches 1102;  Conservative 0;  Mismatches 0;  Indels 1;  Gaps 1;

Qy        1  GAGCTATGCCTACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1224  GAGCTATGCCTACACACAGGCTGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATT 1283

Qy       61  TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1284  TCCTTCACAGCATTTGGAAGTCCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAGAG 1343

Qy      121  TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1344  TGAAGTAAACCTGGACCGTTATCAAACAGCTTAGAAGAAGTATTATCGTGGCTTCTTTC 1403

Qy      181  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAATTCAAGATGGGAAGTGGTGAAAGA 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1404  TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAAATTCAAGATGGGAAGTGGTGAAAGA 1463

Qy      241  CCAGTTCATACTCATGAGGGGTACATGATGAGTTGACAGCCCATCAGGGCCGGGTTGG 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1464  CCAGTTCATACTCATGAGGGGTACATGATGAGTTGACAGCCCATCAGGGCCGGGTTGG 1523

Qy      301  TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1524  TAATATTCTACAATTGGAAGTAAGCTGATTGGAACAGGAAAAATTATCAGAAGATGAAGA 1583

Qy      361  AACTGAAGTACAAGAGACAGAGATGAATCTCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1584  AACTGAAGTACAAGAGACAGAGATGAATCTCCTAAATTCAAGATGGGAAATGCCTCAGGGTAGC 1643

Qy      421  TAGCATGGAAAAACAAAGCAATTACAGAGTTTAATGATCTCCAGAATC-GAAACT 479
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     1644  TAGCATGGAAAAACAAAGCAATTACATAGAGTTTTAATGATCTCCAGAATCAGAAACT 1703
```

```
Qy      480  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAAGAACAAGGAAAATGGAGGAAGA 539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1704  GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAGAAGAAGAACAAGGAAAATGGAGGAAGA 1763

Qy      540  GCCTCTTGACCTGATCTTGAGACTTGAGAGACCTAAAACGCCAGTACAACACATAAGGTGCTTCA 599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1764  GCCTCTTGACCTGATCTTGAGACTTGAGAGACCTAAAACGCCAGTACAACACATAAGGTGCTTCA 1823

Qy      600  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT 659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1824  AGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCTCACTCACATGGTGGTGGTAGT 1883

Qy      660  TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGACAACTTAAGGTATTGGG 719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1884  TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGACAACTTAAGGTATTGGG 1943

Qy      720  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACAT 779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1944  AGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCCGGGTTCTTTTACAAGACAT 2003

Qy      780  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTGTAGTGCATGGCTTTCAGA 839
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2004  CCTTCTCAAATGGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGA 2063

Qy      840  AAAGAAGATGCAGTGAACAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 899
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2064  AAAGAAGATGCAGTGAACAGATTCACACAACTGGCTTAAAGATCAAAATGAAATGTT 2123

Qy      900  ATCAAGTCTTCAAAAACTGGGCCGTTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2124  ATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2183

Qy      960  GGGCAAACTGTATTCACTCAAACAAGATCTCTTTCCCGGTGTTGGGATAATTTAGTCAGTGAC 1019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2184  GGGCAAACTGTATTCACTCAAACAAGATCTCTTTCCCGGTGTTGGGATAATTTAGTCAGTGAC 2243

Qy     1020  CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 1079
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2244  CCAGAAGACGGAGCATGGCTGGATAACTTTGCCCGTGTTGGGATAATTTAGTCCAAAA 2303

Qy     1080  ACTTGAAAAGAGTACAGCACAGA 1102
             |||||||||||||||||||||||
Db     2304  ACTTGAAAAGAGTACAGCACAGA 2326
```

Search completed: April 5, 2004, 16:34:31
Job time : 393.214 secs

```
           Copyright   GenCore version 5.1.6
                    (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on:    April  5, 2004, 04:53:30 ; Search time 315.469 Seconds
                                (without alignments)
                                15454.047 Million cell updates/sec

Title:     US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence:  1 cgactttccagcagttcaga.........ccggagctgcttccaattg 1301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:  2470430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters:   4940860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Published Applications NA:*
         1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
         2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
         3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
         4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
         5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
         6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
         7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
         8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
         9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
        10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
        11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
        12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
        13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
        14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
        15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
        16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
        17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
        18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1301	100.0	1301	1821	10	US-09-845-416-13	Sequence 13, Appl
2	1301	100.0	1301	2169	10	US-09-845-416-4	Sequence 4, Appli
3	1301	100.0	1301	3510	10	US-09-845-416-12	Sequence 12, Appl
4	1301	100.0	1301	3531	10	US-09-845-416-10	Sequence 10, Appl
5	1301	100.0	1301	3858	10	US-09-845-416-9	Sequence 9, Appli
6	1301	100.0	1301	3999	10	US-09-845-416-6	Sequence 6, Appli
7	1301	100.0	1301	4182	10	US-09-845-416-2	Sequence 2, Appli
8	1301	100.0	1301	4476	10	US-09-845-416-31	Sequence 31, Appl
9	1301	100.0	1301	4498	10	US-09-845-416-30	Sequence 30, Appl
10	1301	100.0	1301	4825	10	US-09-845-416-29	Sequence 29, Appl
11	1301	100.0	1301	4848	10	US-09-845-416-35	Sequence 35, Appl
12	1301	100.0	1301	4966	10	US-09-845-416-28	Sequence 28, Appl
13	1301	100.0	1301	4990	10	US-09-845-416-34	Sequence 34, Appl
14	1301	100.0	1301	5060	10	US-09-845-416-36	Sequence 36, Appl
15	1301	100.0	1301	5149	10	US-09-845-416-27	Sequence 27, Appl
16	1301	100.0	1301	5339	15	US-10-149-736-40	Sequence 40, Appl
17	1301	100.0	1301	5462	15	US-10-149-736-41	Sequence 41, Appl
18	1301	100.0	1301	8689	15	US-10-149-736-42	Sequence 42, Appl
19	1301	100.0	1301	11058	15	US-09-845-416-1	Sequence 1, Appli
20	1301	100.0	1301	11443	15	US-10-149-736-44	Sequence 44, Appl
21	1301	100.0	1301	12057	15	US-10-149-736-47	Sequence 47, Appl
22	1301	100.0	1301	13957	9	US-09-782-378A-22	Sequence 22, Appl
23	1301	100.0	1301	13957	9	US-09-880-107-2284	Sequence 2284, Ap
24	1301	100.0	1301	13957	15	US-10-149-736-1	Sequence 1, Appli
25	1301	100.0	1301	14069	12	US-10-342-887-434	Sequence 434, App
26	1301	100.0	1301	14082	12	US-10-342-887-981	Sequence 981, App
27	1301	100.0	1301	14082	15	US-10-341-434-108	Sequence 108, App
28	1135.2	87.3	1301	13815	15	US-10-149-736-2	Sequence 2, Appli
29	1017.6	78.2	1301	3446	10	US-09-845-416-14	Sequence 14, Appl
30	1017.6	78.2	1301	4414	10	US-09-845-416-32	Sequence 32, Appl
31	1016.6	78.1	1301	5417	10	US-09-845-416-15	Sequence 15, Appl
32	1016	78.1	1301	1434	10	US-09-845-416-15	Sequence 15, Appl
33	567.2	43.6	1301	11096	15	US-10-149-736-4	Sequence 4, Appli
34	553.8	42.6	1301	10302	9	US-09-782-378A-23	Sequence 23, Appl
35	553.8	42.6	1301	10302	15	US-10-149-736-3	Sequence 3, Appli
36	551	42.4	1301	5106	15	US-10-220-120-157	Sequence 157, App
37	545	41.9	1301	16531	14	US-10-101-510-667	Sequence 667, App
38	472	36.3	1301	887	15	US-10-149-736-35	Sequence 35, Appl
39	324	24.9	1301	324	15	US-10-149-736-33	Sequence 33, Appl
40	289	22.2	1301	387	15	US-10-149-736-32	Sequence 32, Appl
41	216	16.6	1301	216	15	US-10-149-736-34	Sequence 34, Appl
42	114	8.8	1301	114	15	US-10-149-736-45	Sequence 45, Appl
c 43	102.8	7.9	1301	275	14	US-10-029-386-16813	Sequence 16813, A
c 44	102.8	7.9	1301	587	14	US-10-029-386-3113	Sequence 3113, Ap
45	54.2	4.2	1301	449	10	US-09-918-995-24084	Sequence 24084, A

ALIGNMENTS

```
RESULT 1
US-09-845-416-13
; Sequence 13, Application US/0984516
; Publication No. US2003017131241
; GENERAL INFORMATION:
;   APPLICANT: XIAO, XIAO
;   TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;   TITLE OF INVENTION: THEREOF
;   FILE REFERENCE: DE1142
;   CURRENT APPLICATION NUMBER: US/09/845,416
;   CURRENT FILING DATE: 2001-04-30
;   PRIOR APPLICATION NUMBER: 60/200,777
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 13
;   LENGTH: 1821
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-845-416-13

Query Match             100.0%;  Score 1301;  DB 10;  Length 1821;
Best Local Similarity   100.0%;
Matches 1301;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy          1 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGGGGAATTGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGGGGAATTGAA 162

Qy         61 AACTAAAGAACCTGTAATCATGATGACTACTTCTGAGACTGTACGAAATATTTCTGACAGAGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AACTAAAGAACCTGTAATCATGATGACTACTTCTGAGACTGTACGAAATATTTCTGACAGAGCA 222

Qy        121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTCCTCCTGAGGAGAG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTCGCTCCTGAGGAGAG 282

Qy        181 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTACGAAAAGCAGAGGCGGTGAGGAGTCAATACTGAGTGGGA 240
```

```
Db:  1363 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1403
```

RESULT 2
US-09-845-416-4
; Sequence 4, Application US/0984 5416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-4

Query Match   100.0%; Score 1301; DB 10; Length 2169;
Best Local Similarity   100.0%;
Matches 1301; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```

```
Qy   1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 451 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAA 510

Qy   61 AACTAAAGAACCTGTAATCATGAGTACTCTGAGACTGTACGAGAATATTTCTGACAGAGCA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 511 AACTAAAGAACCTGTAATCATGAGTACTCTGAGACTGTACGAGAATATTTCTGACAGAGCA 570

Qy 121 GCCTTTGAAGGACTAGAGAAACTTCTACCAGGAGAGCTGCCTCCTGAGGAGAG   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 571 GCCTTTGAAGGACTAGAGAAACTTCTACCAGGAGAGCTGCCTCCTGAGGAGAG   630

Qy 181 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 631 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA   690

Qy 241 AAAATTGAACCTGCACTCGGCTGACTCGGCTGCAGAGAAAATAGATGAGACCCTTGAAAGACT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 691 AAAATTGAACCTGCACTCGGCTGACTCGGCTGCAGAGAAAATAGATGAGACCCTTGAAAGACT 750

Qy 301 CCAGGAACTTCAAGAGAGCCACGATGAGCTGACTCGGCCAAGCTGCGGCCAAGCTGAGGTGAT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 751 CCAGGAACTTCAAGAGAGCCACGATGAGCTGACTCGGCCAAGCTGCGGCCAAGCTGAGGTGAT 810

Qy 361 CAAGGGATCCTGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 811 CAAGGGATCCTGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA   870

Qy 421 GAAAGTCAAGGCACTTGGAGAGAGAAATTGCGGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 871 GAAAGTCAAGGCACTTGGAGAGAGAAATTGCGGCCTCTGAAAGAGAACGTGAGCCACGTCAA 930

Qy 481 TGACCTTGCTCGCCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 931 TGACCTTGCTCGCCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 990

Qy 541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTCGCAGGTGGCCGTCGCAGGACCGAGTCAG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 991 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTCGCAGGTGGCCGTCGCAGGACCGAGTCAG 1050

Qy 601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGACTTTCTTCCACGTC   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1051 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGACTTTCTTCCACGTC   1110

Qy 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA   720
```

```
Db  283 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA   342
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 AAAATTGAACCTGCACTCGGCTGACTCGGCTGCAGAGAAAATAGATGAGACCCTTGAAAGACT 300

Db  343 AAAATTGAACCTGCACTCGGCTGACTCGGCTGCAGAGAAAATAGATGAGACCCTTGAAAGACT 402
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 CCAGGAACTTCAAGAGAGCCACGATGAGCTGACTCGGCCAAGCTGCGGCCAAGCTGAGGTGAT 360

Db  403 CCAGGAACTTCAAGAGAGCCACGATGAGCTGACTCGGCCAAGCTGCGGCCAAGCTGAGGTGAT 462
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 CAAGGGATCCTGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA    420

Db  463 CAAGGGATCCTGCAGCCCGTGGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA    522
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421 GAAAGTCAAGGCACTTGGAGAGAGAAATTGCGGCCTCTGAAAGAGAACGTGAGCCACGTCAA  480

Db  523 GAAAGTCAAGGCACTTGGAGAGAGAAATTGCGGCCTCTGAAAGAGAACGTGAGCCACGTCAA  582
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  481 TGACCTTGCTCGCCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540

Db  583 TGACCTTGCTCGCCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 642
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTCGCAGGTGGCCGTCGCAGGACCGAGTCAG  600

Db  643 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTCGCAGGTGGCCGTCGCAGGACCGAGTCAG  702
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGACTTTCTTCCACGTC       660

Db  703 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGACTTTCTTCCACGTC       762
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA   720

Db  763 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA   822
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721 CGAGACTCAAAACAACTTGCTCGGGACCATCCCAAAATGACAGAGCTCTTTAGC        780

Db  823 CGAGACTCAAAACAACTTGCTCGGGACCATCCCAAAATGACAGAGCTCTTTAGC        882
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  781 TGACCTGAATAATGTCAGATTCTCAGCTATAGGACTGCCATGAAACTCCGAAGACTGCA   840

Db  883 TGACCTGAATAATGTCAGATTCTCAGCTATAGGACTGCCATGAAACTCCGAAGACTGCA   942
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  841 GAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCATGTGGACCAGCA  900

Db  943 GAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCATGTGGACCAGCA 1002
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901 CAACCTCAAGACAAATGACCAGCCCATGGAGATATCCTGCAGATTAATTGTTTGACCAC    960

Db 1003 CAACCTCAAGACAAATGACCAGCCCATGGAGATATCCTGCAGATTAATTGTTTGACCAC   1062
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961 TATTTATGACCGGCCTGGAGCAAGAGCAACAACAATTGGTCAACGTCCTCGGCTGGA    1020

Db 1063 TATTTATGACCGGCCTGGAGCAAGAGCAACAACAATTGGTCAACGTCCTCGGCTGGA    1122
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACAGGGAGGGATCCGTGT   1080

Db 1123 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACAGGGAGGGATCCGTGT   1182
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081 CCTGTCTTTTAAAACTTCAAGCAAGTGGCATCATTCCCTGTGTAAAGCAAGTTGGAAGACAAGTACAG 1140

Db 1183 CCTGTCTTTTAAAACTTCAAGCAAGTGGCATCATTCCCTGTGTAAAGCAAGTTGGAAGACAAGTACAG 1242
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCATCCAAAATTCCAAGACCAGTGGGTGAGT         1200

Db 1243 ATACCTTTTCAAGCAAGTGGCAAGTTCATCCAAAATTCCAAGACCAGTGGGTGAGT         1302
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1201 CCTTCTGCATGATTCTATCCAAGAAGACACAGGACTTTGGTGAAGTTGCATCCTTTGGGGG   1260

Db 1303 CCTTCTGCATGATTCTATCCAAGAAGACACAGGACTTTGGTGAAGTTGCATCCTTTGGGGG   1362
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG                     1301
```

```
Qy    61  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTCTGACAGAGCA   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1839  AACTAAAGAACCTGTAATCA'GAGTACTCA'GAGTACTCTTGAGACTGTACGAGAATATTCTGACAGAGCA  1898

Qy   121  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGAAACTCTACCAGGAGAGCTGCCTCCTGAGGAGAG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1899  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGAAACTCTACCAGGAGAGCTGCCTCCTGAGGAGAG  1958

Qy   181  AGCCCAGAAGTCACTCGGCTTCAGCTGCCATGAAACTCCGAAGACTGCA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1959  AGCCCAGAAGTCACTCGGCTTCAGCTGCCATGAAACTCCGAAGACTGCA  2018

Qy   241  AAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2019  AAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT  2078

Qy   301  CCAGGAACTTCAAGAGCTCAGCGATGAGCTGACCTCAAGTGCGCCAAGTGAGGTGAT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2079  CCAGGAACTTCAAGAGCTCAGCGATGAGCTGACCTCAAGTGCGCCAAGTGAGGTGAT  2138

Qy   361  CAAGGGATCCTGCAGCCCGTGGCGATCTCCTCATTGACTCTCCAAGATCACCTCGA   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2139  CAAGGGATCCTGCAGCCCGTGGCGATCTCCTCATTGACTCTCCAAGATCACCTCGA  2198

Qy   421  GAAAGTCAAGGACACTTGAGAGGAGAAATTGCGCCTTGAAAGAGAACGTGAGCCAGTCAA   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2199  GAAAGTCAAGGACACTTGAGAGGAGAAATTGCGCCTTGAAAGAGAACGTGAGCCAGTCAA  2258

Qy   481  TGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTCAGTGGCGTCTCAGCACCTAGCAC   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2259  TGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTCAGTGGCGTCTCAGCACCTAGCAC  2318

Qy   541  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGTGGCGTCGAGACCGAGTCAG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2319  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGTGGCGTCGAGACCGAGTCAG  2378

Qy   601  GCAGCTGCATGAGCCTCCAGGGACTTTGGTCCAGATCTCAGCACTTCTTTCCACGTC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2379  GCAGCTGCATGAGCCTCCAGGGACTTTGGTCCAGATCTCAGCACTTCTTTCCACGTC  2438

Qy   661  TGTCCAGGTGTCCCTGGGAGGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2439  TGTCCAGGTGTCCCTGGGAGGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA  2498

Qy   721  CGAGACTCAAACAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2499  CGAGACTCAAACAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC  2558

Qy   781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2559  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  2618

Qy   841  GAAGGCCCTTTGCTTGGATCTCTTGGATCTCTTGAGCGTGCAGTGCATGTGATGCCTTGGACCAGCA   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2619  GAAGGCCCTTTGCTTGGATCTCTTGGATCTCTTGAGCGTGCAGTGCATGTGATGCCTTGGACCAGCA  2678

Qy   901  CAACCTCAAGCAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2679  CAACCTCAAGCAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC  2738

Qy   961  TATTTATGACCGCCTGGAACTGGCTGCTGAATGTTATGATACGGGAACATTTGGAAGACAAGTACAG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2739  TATTTATGACCGCCTGGAACTGGCTGCTGAATGTTATGATACGGGAACATTTGGAAGACAAGTACAG  2798

Qy  1021  TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGAGATCCGTGT  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2799  TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGAGATCCGTGT  2858

Qy  1081  CCTGTCTTTTAAAACTGGACATCATTTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2859  CCTGTCTTTTAAAACTGGCATCATTTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG  2918

Qy  1141  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGTTCAAGAGAGGGAATTGAA  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1779  CGACTTCCAGCAGTTCAAGCAGAACGATGCACATAGGGCCTTCAAGAGGGAATTGAA  1838
```

```
Db  1111  TGTCCAGGTGTCCCTGGGAGGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA  1170
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721  CGAGACTCAAACAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC   780

Db  1171  CGAGACTCAAACAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC  1230
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA   840

Db  1231  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  1290
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841  GAAGGCCCTTTGCTTGGATCTCTTGAGCGTCAGTGCATGTGATGCCTTGGACCAGCA   900

Db  1291  GAAGGCCCTTTGCTTGGATCTCTTGAGCGTCAGTGCATGTGATGCCTTGGACCAGCA  1350
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901  CAACCTCAAGCAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC   960

Db  1351  CAACCTCAAGCAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC  1410
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   961  TATTTATGACCGCCTGGAACTGGCTGCTGAATGTTATGATACGGGAACATTTGGAAGACAAGTACAG  1020

Db  1411  TATTTATGACCGCCTGGAACTGGCTGCTGAATGTTATGATACGGGAACATTTGGAAGACAAGTACAG  1470
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1021  TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGAGATCCGTGT  1080

Db  1471  TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGAGATCCGTGT  1530
          ||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1081  CCTGTCTTTTAAAACTGGCATCATTTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG  1140

Db  1531  CCTGTCTTTTAAAACTGGCATCATTTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG  1590
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1141  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGTTCAAGAGAGGGAATTG  1200

Db  1591  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGTTCAAGAGAGGGAATTG  1650
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1201  CCTTCTGCATGATTCTATCCAAGCACAGTTGGGTGGTGAAGTTGCATCCTTTGGGGG  1260

Db  1651  CCTTCTGCATGATTCTATCCAAGCACAGTTGGGTGGTGAAGTTGCATCCTTTGGGGG  1710
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301

Db  1711  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1751
```

RESULT 3
US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US2003017131 2A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match   100.0%; Score 1301; DB 10; Length 3510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1  CGACTTCCAGCAGTTCAAGCAGAACGATGCACATAGGGCCTTCAAGAGGGAATTGAA    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1779  CGACTTCCAGCAGTTCAAGCAGAACGATGCACATAGGGCCTTCAAGAGGGAATTGAA  1838
```

```
Qy  541  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGCCGTCGAGGACCGAGTCAG   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2340  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGCCGTCGAGGACCGAGTCAG  2399

Qy  601  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGCACTTCTTCCCACGTC    660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2400  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGCACTTCTTCCCACGTC   2459

Qy  661  TGTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCA   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2460  TGTCCAGGGTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCA  2519

Qy  721  CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCACCAGTCTTTAGC    780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2520  CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCACCAGTCTTTAGC   2579

Qy  781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTTATAGGACTCCGAAGACTGCA    840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2580  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTTATAGGACTCCGAAGACTGCA   2639

Qy  841  GAAGGCCCTTGCTTGGATCTCTTGAGCGTGCAGCTGCAGTGATGCCTTGGACCAGCA    900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2640  GAAGGCCCTTGCTTGGATCTCTTGAGCGTGCAGCTGCAGTGATGCCTTGGACCAGCA   2699

Qy  901  CAACCTCAAGCAACAAATGACCAGCCATGGATATCCTGCAGATTATTAATTGTTGACCAC   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2700  CAACCTCAAGCAACAAATGACCAGCCATGGATATCCTGCAGATTATTAATTGTTGACCAC  2759

Qy  961  TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCCTCGCGTGGA  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2760  TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCCTCGCGTGGA  2819

Qy 1021  TATGTGTCTGAACTGGCTGCTGTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2820  TATGTGTCTGAACTGGCTGCTGTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT  2879

Qy 1081  CCTGTCTTTAAAACTGGCATCATTCCCTGTAAACACATTTGGAAGACAAGTACAG  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2880  CCTGTCTTTAAAACTGGCATCATTCCCTGTAAACACATTTGGAAGACAAGTACAG  2939

Qy 1141  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGGCTGGGCCT  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2940  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGGCTGGGCCT  2999

Qy 1201  CCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAGTTGCATCCTTCAAATTG  1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3000  CCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAGTTGCATCCTTCAAATTG  3059

Qy 1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301
         |||||||||||||||||||||||||||||||||||||||||
Db 3060  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  3100
```

```
Db 2919  ATACCTTTTCAAGCAAGTGGCAAGTTCATCCAAATTCCAAGACAGATTTGTGACCAGCGCAGGCTGGGCCT  2978
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1201  CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGCTTCCGGAGCTGCTTCCAATTTG  1260
Db 2979  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  3038
Qy 1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301
Db 3039  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  3079
```

```
Query Match            100.0%;  Score 1301;   DB 10;  Length 3531;
Best Local Similarity  100.0%;
Matches 1301;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1  CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1800  CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA  1859

Qy   61  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1860  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA  1919

Qy  121  GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCGCCTCCTCCTCCTGAGGAGGAG   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1920  GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCGCCTCCTCCTCCTGAGGAGGAG  1979

Qy  181  AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCAGGAGCAGCAGGAGCTGAGTGGGA   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1980  AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCAGGAGCAGCAGGAGCTGAGTGGGA  2039

Qy  241  AAAATGAACCTGCACTCCGCTGACTGGCAGCAGAGAAAATAGATGAGACCCTTGAAAGACT   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2040  AAAATTGAACCTGCACTCCGCTGACTGGCAGCAGAGAAAAATAGATGAGACCCTTGAAAGACT  2099

Qy  301  CCAGGAACTTCAAGAGAGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGAT   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2100  CCAGGAACTTCAAGAGAGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGGTGAT  2159

Qy  361  CAAGGGATCCTGCGCAGCCCGTGTGGGCCATCTCCTCATTGACTCCTCCAAGATCACCTCGA   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2160  CAAGGGATCCTGCGCAGCCCGTGTGGGCCATCTCCTCATTGACTCCTCCAAGATCACCTCGA  2219

Qy  421  GAAAGTCAAGGACTTCAGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAA   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2220  GAAAGTCAAGGACTTCAGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAA  2279

Qy  481  TGACCTTGCTCGCCAGCTTACCCACTTTGGGCCATTCAGCTCCAGCTCTCAGCAC   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2280  TGACCTTGCTCGCCAGCTTACCCACTTTGGGCCATTCAGCTCCAGCTCTCAGCAC  2339
```

Query Match   100.0%; Score 1301; DB 10; Length 3858;
Best Local Similarity 100.0%; Pred. No.0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2127 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2186

Qy    61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGATATTTCTGACAGAGCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2187 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGATATTTCTGACAGAGCA 2246

Qy   121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2247 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 2306

Qy   181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2307 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2366

Qy   241 AAAATTGAACCTGCACTCGCTGACTCGGCAGCAGAAAAATAGATGAGACCCTTGAAAGACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2367 AAAATTGAACCTGCACTCGCTGACTCGGCAGCAGAAAAATAGATGAGACCCTTGAAAGACT 2426

Qy   301 CCAGGAACTTCAAGAGGCCACGATGAGCTGGACCTGCAAGCTGCGCCAAGCTGAGGTGAT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2427 CCAGGAACTTCAAGAGGCCACGATGAGCTGGACCTGCAAGCTGCGCCAAGCTGAGGTGAT 2486

Qy   361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTTCCTCTTCCAAGATCACCTCGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2487 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTTCCTCTTCCAAGATCACCTCGA 2546

Qy   421 GAAAGTCAAGGCACTTCGACGAGAAATTGCGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2547 GAAAGTCAAGGCACTTCGACGAGAAATTGCGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 2606

Qy   481 TGACCTGCTGCCGCTTACCACCTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2607 TGACCTGCTGCCGCTTACCACCTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 2666

Qy   541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCCAGGTGGCCGTCGAGGACCGAGTCAG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2667 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCCAGGTGGCCGTCGAGGACCGAGTCAG 2726

Qy   601 GCAGCTGCATGAAGCCCACAGGACTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2727 GCAGCTGCATGAAGCCCACAGGACTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 2786

Qy   661 TGTCCAGGGTCCTGGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2787 TGTCCAGGGTCCTGGAGAGCCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 2846

Qy   721 CGAGACTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2847 CGAGACTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGCTCTTACCAGTCTTTAGC 2906

Qy   781 TGACCTGAATAATGTCAGATTCTCAGCTTCTGAGCTGCCATGAAACTCCGAAGACTGCA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2907 TGACCTGAATATGTCAGATTCTCAGCTTCTGAGCTGCCATGAAACTCCGAAGACTGCA 2966

Qy   841 GAAGGGCCCTTTGCTTGGATCTCGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2967 GAAGGGCCCTTTGCTTGGATCTCGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA 3026

Qy   901 CAACCTTCAAGCAAAATGACCAGCCCATGGATATCCCAGATTATTAATTGTTGACCAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3027 CAACCTTCAAGCAAAATGACCAGCCCATGGATATCCCAGATTATTAATTGTTGACCAC 3086

Qy   961 TATTTATGACCGCCTGAGCAAGAGCAAGAGCACAACAATTTGGTCAAGTCCCTCTCGCGTGA 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3087 TATTTATGACCGCCTGAGCAAGAGCAAGAGCACAACAATTTGGTCAAGTCCCTCTCGCGTGA 3146
```

```
Qy  1021 TATGTGTCTGAACTGGCTCTGAATGTTTATGATACGGGACGAACAGGGAGGAGGATCCGTGT 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3147 TATGTGTCTGAACTGGCTCTGAATGTTTATGATACGGGACGAACAGGGAGGAGGATCCGTGT 3206

Qy  1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3207 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 3266

Qy  1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGGCAGGCTGGGCCT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3267 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGGCAGGCTGGGCCT 3326

Qy  1201 CCTTCTGCGATGATCTATCCAAATTCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3327 CCTTCTGCGATGATCTATCCAAATTCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 3386

Qy  1261 CAGTAACACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
           |||||||||||||||||||||||||||||||||||||||||||
Db  3387 CAGTAACACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 3427
```

RESULT 6
US-09-845-416-6

; Sequence 6, Application US/09845416
; Publication No. US2003017l312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; SOFTWARE: PatentIn Ver. 2.1
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6

Query Match   100.0%; Score 1301; DB 10; Length 3999;
Best Local Similarity 100.0%; Pred. No.0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2268 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2327

Qy    61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGATATTTCTGACAGAGCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2328 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGATATTTCTGACAGAGCA 2387

Qy   121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2388 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 2447

Qy   181 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2448 AGCCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2507

Qy   241 AAAATTGAACCTGCACTCGCTGACTCGGCAGCAGAAAAATAGATGAGACCCTTGAAAGACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2508 AAAATTGAACCTGCACTCGCTGACTCGGCAGCAGAAAAATAGATGAGACCCTTGAAAGACT 2567

Qy   301 CCAGGAACTTCAAGAGGCCACGATGAGCTGGACCTGCAAGCTGCGCCAAGCTGAGGTGAT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2568 CCAGGAACTTCAAGAGGCCACGATGAGCTGGACCTGCAAGCTGCGCCAAGCTGAGGTGAT 2627

Qy   361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTTCCTCTTCCAAGATCACCTCGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2628 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTTCCTCTTCCAAGATCACCTCGA 2687
```

Query Match   100.0%; Score 1301; DB 10; Length 4182;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGCAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2451 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGCAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2510

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCCTTGAGACTGTACAGAGCA                      120
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2511 AACTAAAGAACCTGTAATCATGAGTACTCCTTGAGACTGTACAGAGCA                      2570

Qy    121 GCCTTGAAGACTAGAGAAACTCTACCAGGAGAGCTGCCTCCTGAGGAGAG                     180
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2571 GCCTTGAAGACTAGAGAAACTCTACCAGGAGAGCTGCCTCCTGAGGAGAG                    2630

Qy    181 AGCCCAGAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA            240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2631 AGCCCAGAATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA            2690

Qy    241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT           300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2691 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT           2750

Qy    301 CCAGGAACTTCAAGAGCCACGGATGAGCTGACCTCCTCATTGACTCTCTCCAGATCACCTCGA        360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2751 CCAGGAACTTCAAGAGCCACGGATGAGCTGACCTCCTCATTGACTCTCTCCAGATCACCTCGA        2810

Qy    361 CAAGGGATCCTGGCGCAGCCCGTGGGCTCGGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAA  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2811 CAAGGGATCCTGGCGCAGCCCGTGGGCTCGGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAA  2870

Qy    421 GAAAGTCAAGGCACTTGGTGGCCAGCTTACCACTTTGGTCATGCACATTTGGAAGACAAGTACAG       480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2871 GAAAGTCAAGGCACTTGGTGGCCAGCTTACCACTTTGGTCATGCACATTTGGAAGACAAGTACAG       2930

Qy    481 TGACCTTGCTCGGCAGCTTACCACTTTGGGCATTCAGCTCTCCACGGTATAACCTCAGCAG         540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2931 TGACCTTGCTCGGCAGCTTACCACTTTGGGCATTCAGCTCTCCACGGTATAACCTCAGCAC         2990

Qy    541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGGAGGACCGAGTCAG          600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2991 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGGAGGACCGAGTCAG          3050

Qy    601 GCAGCTGCATGAAGCCCAGATCATCTGCGCAAACAAAATGACAGAGCTTCTTACCAGTCTTTAGC      660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 GCAGCTGCATGAAGCCCAGATCATCTGCGCAAACAAAATGACAGAGCTTCTTACCAGTCTTTAGC      3110

Qy    661 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGCAAACTCCGAAGACTGCA          720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGCAAACTCCGAAGACTGCA          3170

Qy    721 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTTCTACCAGTCTTTAGC         780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3171 CGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTTCTACCAGTCTTTAGC          3230

Qy    781 GAAGGCCCTTGCTTGTCAGTCTCGTCAGCTGCAGTTTATGGATAATGTCAGATTTTAAAC          840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 GAAGGCCCTTGCTTGTCAGTCTCGTCAGCTGCAGTTTATGGATAATGTCAGATTTTAAAC          3290

Qy    841 CAACTCAAGCAGCAAATGACCAGCCCATGAATATCCTGCAGATATTAATTGTTTGACCAC          900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 GAAGGCCCTTGCTTGTCAGTCTCGTCAGCTGCAGTTTATGGATAATGTCAGATTTTAAAC          3350
```

[DNA sequence alignment blocks — Qy/Db paired sequences with match bars, positions 2775–3835 / 301–1301]

RESULT 8
US-09-845-416-31
Sequence 31, Application US/09845416
Publication No. US2003017131A1
GENERAL INFORMATION:
  APPLICANT: XIAO, XIAO
  TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
  TITLE OF INVENTION: THEREOF
  FILE REFERENCE: DE1142
  CURRENT APPLICATION NUMBER: US/09/845,416
  CURRENT APPLICATION FILING DATE: 2001-04-30
  PRIOR APPLICATION NUMBER: 60/200,777
  PRIOR APPLICATION FILING DATE: 2000-04-28
  SOFTWARE: PatentIn Ver.2.1
  NUMBER OF SEQ ID NOS: 36
  SEQ ID NO 31
  LENGTH: 4476
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-845-416-31

Query Match   100.0%; Score 1301; DB 10; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[DNA sequence alignment blocks — Qy/Db paired sequences with match bars, positions 1–300 / 2535–2774]

RESULT 9
US-09-845-416-30

; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-30

Query Match   100.0%;   Score 1301;   DB 10;   Length 4498;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1301;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2557 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2616

Qy   61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2617 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2676

Qy 121 GCCTTTGGAAGGACTAGAGAAAACTTCACCAGGAGGACCCAGAGAGCTGCTTCCTCTGAGGAGAG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2677 GCCTTTGGAAGGACTAGAGAAAACTTCACCAGGAGGACCCAGAGAGCTGCTTCCTCTGAGGAGAG 2736

Qy 181 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2737 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2796

Qy 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2797 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAGAAAAATAGATGAGACCCTTGAAAGACT 2856

Qy 301 CCAGGAACTTCAAGAGCTTCAAGAGGCCACGGATGAGCTGGACTTGGAACTGCACTCGCGCAAGCTGCGCCAAGCTGAGGTGAT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2857 CCAGGAACTTCAAGAGCTTCAAGAGGCCACGGATGAGCTGGACTTGGAACTGCACTCGCGCAAGCTGCGCCAAGCTGAGGTGAT 2916

Qy 361 CAAGGGATCCTCGGCGAGCCCGTGGGCAGTCTCCCATTGACTCCTCCAAGATCACTCGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2917 CAAGGGATCCTCGGCGAGCCCGTGGGCAGTCTCCCATTGACTCCTCCAAGATCACTCGA 2976

Qy 421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTGGGAATGTGAGCCACGTCAA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2977 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTTGGGAATGTGAGCCACGTCAA 3036

Qy 481 TGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCACGTATAACCTCAGCAC 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3037 TGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCACGTATAACCTCAGCAC 3096

Qy 541 TCTGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCAGGTGGCCGTCGAGGACCGAGTCAG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3097 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCGTCAGGTGGCCGTCGAGGACCGAGTCAG 3156

Qy 601 GCAGCTGCATGAAGCTGAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3157 GCAGCTGCATGAAGCTGAGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3216

Qy 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3217 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3276

Qy 721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 3277 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3336

Qy 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3337 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3396

Qy 841 GAAGGCCCTTGCTTGGACTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGACCAGCA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3397 GAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGACCAGCA 3456

Qy 901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3457 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 3516

Qy 961 TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGTCACAACGTCCTCTCTGCGTGGA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3517 TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGTCACAACGTCCTCTCTGCGTGGA 3576

Qy 1021 TATGTGTCTGAACTGGCTCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3577 TATGTGTCTGAACTGGCTCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 3636

Qy 1081 CCTGTCTTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3637 CCTGTCTTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG 3696

Qy 1141 ATACCTTTCAAGCAAGTGCAAGTTCAACAGGATTTTGTGACCAGCCAGGCTGGGCCT 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3697 ATACCTTTCAAGCAAGTGCAAGTTCAACAGGATTTTGTGACCAGCCAGGCTGGGCCT 3756

Qy 1201 CCTTCTGCATGATTCTATCCAAATTCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3757 CCTTCTGCATGATTCTATCCAAATTCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 3816

Qy 1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
         |||||||||||||||||||||||||||||||||||||||||
Db 3817 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 3857

RESULT 10
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match   100.0%;   Score 1301;   DB 10;   Length 4825;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1301;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2884 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2943

Qy   61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2944 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3003

Qy 121 GCCTTTGGAAGGACTAGAGAAAACTTCACCAGGAGGACCCAGAGAGCTGCCTCCTGAGGAGAG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db  4084  CCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG  4143
              |||||||||||||||||||||||||||||||||||||||||
Qy  1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301
              |||||||||||||||||||||||||||||||||||||||||
Db  4144  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  4184


RESULT 11
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
;   APPLICANT: XIAO, XIAO
;   TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
;   TITLE OF INVENTION:  THEREOF
;   FILE REFERENCE: DE1142
;   CURRENT APPLICATION NUMBER: US/09/845,416
;   CURRENT FILING DATE: 2001-04-30
;   PRIOR APPLICATION NUMBER: 60/200,777
;   PRIOR FILING DATE: 2000-04-28
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
;   LENGTH: 4848
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-845-416-35

Query Match            100.0%;   Score 1301;   DB 10;   Length 4848;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 1301;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy     1  CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2907  CGACTTTCCAGCAGTTCAGAAGCAGAACGATGATGTACATAGGGCCTTCAAGAGGGAATTGAA  2966

Qy    61  AACTAAAGAACCTGTAATCATGAGTACTCCTTGAGACTGTACGACAGAGCA   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2967  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGACAGAGCA   3026

Qy   121  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCAGCCCAGAGAGCTGCCTCCTGAGGAGAG   180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3027  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCAGCCCAGAGAGCTGCCTCCTGAGGAGAG   3086

Qy   181  AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTCACGAAAGCAGAGGCTCAATACTGAGTGGGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3087  AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTCACGAAAGCAGAGGCTCAATACTGAGTGGGA   3146

Qy   241  AAAATTGAACCTGCACTCCGTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3147  AAAATTGAACCTGCACTCCGTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT   3206

Qy   301  CCAGGAACTTCAAGAGGCCACGCACCGGATGAGCTGACCTCAAGCTGAGGTGAT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3207  CCAGGAACTTCAAGAGGCCACGCACCGGATGAGCTGACCTCAAGCTGAGGTGAT   3266

Qy   361  CAAGGGATCCTGGCAGCCGTGGGCGATCTCAGCTGACTCTCTCAAGATCACCTGA   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3267  CAAGGGATCCTGGCAGCCGTGGGCGATCTCAGCTGACTCTCTCAAGATCACCTGA   3326

Qy   421  GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAA   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3327  GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAA   3386

Qy   481  TGACCTTGCTGCTCGGCCAGCTTACCACTTTGTGACCAGTCTTTAGC   540
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3387  TGACCTTGCTGCTCGGCCAGCTTACCACTTTGTGACCAGTCTTTAGC   3446

Qy   541  TCTTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3447  TCTTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG   3506

Qy   601  GCAGCTGCATGAGCGCATGCTGCAGTTGGTCCAGCAGCACTTTCTTCCACGTC   660
```

```
Db  3004  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCAGCCCAGAGAGCTGCCTCCTGAGGAGAG  3063
Qy   181  AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTCACGAAAGCAGAGGCTCAATACTGAGTGGGA  240
Db  3064  AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTCACGAAAGCAGAGGCTCAATACTGAGTGGGA  3123
Qy   241  AAAATTGAACCTGCACTCCGTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT  300
Db  3124  AAAATTGAACCTGCACTCCGTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT  3183
Qy   301  CCAGGAACTTCAAGAGGCCACGCACCGGATGAGCTGACCTCAAGCTGAGGTGAT  360
Db  3184  CCAGGAACTTCAAGAGGCCACGCACCGGATGAGCTGACCTCAAGCTGAGGTGAT  3243
Qy   361  CAAGGGATCCTGGCAGCCGTGGGCGATCTCCTCAAGATCACCTGA  420
Db  3244  CAAGGGATCCTGGCAGCCGTGGGCGATCTCCTCAAGATCACCTGA  3303
Qy   421  GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAA  480
Db  3304  GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAACGTGAGCCACGTCAA  3363
Qy   481  TGACCTTGCTGCTCGGCCAGCTTACCACTTTGTGACCAGTCTTTAGC  540
Db  3364  TGACCTTGCTGCTCGGCCAGCTTACCACTTTGTGACCAGTCTTTAGC  3423
Qy   541  TCTTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG  600
Db  3424  TCTTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG  3483
Qy   601  GCAGCTGCATGAGCGCATGCTGCAGTTGGTCCAGCAGCACTTTCTTCCACGTC  660
Db  3484  GCAGCTGCATGAGCGCATGCTGCAGTTGGTCCAGCAGCACTTTCTTCCACGTC  3543
Qy   661  TGTCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA  720
Db  3544  TGTCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA  3603
Qy   721  CGAGACTCAAACAACTGCTGGACCATCCAAAATGACACAGAGCTCTACCAGTCTTTAGC  780
Db  3604  CGAGACTCAAACAACTGCTGGACCATCCAAAATGACACAGAGCTCTACCAGTCTTTAGC  3663
Qy   781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAAGACTGCA  840
Db  3664  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAAGACTGCA  3723
Qy   841  GAAGGCCCTTGCTTGGATCCTTGAGCCTGTCAGCTGCATGGAATGTGATGCCTTGGACCAGCA  900
Db  3724  GAAGGCCCTTGCTTGGATCCTTGAGCCTGTCAGCTGCATGGAATGTGATGCCTTGGACCAGCA  3783
Qy   901  CAACCTCAAGCAAAATGACCAGCCATGGAATCCCATGGATATTCCTGCAGATTATTAATTGTTGACCAC  960
Db  3784  CAACCTCAAGCAAAATGACCAGCCATGGAATCCCATGGATATTCCTGCAGATTATTAATTGTTGACCAC  3843
Qy   961  TATTTATGCACGCCTGGAGCAAGAGCAAGAGCACAACAATTGTCACGTCCCTCTCGCGTGGA  1020
Db  3844  TATTTATGCACGCCTGGAGCAAGAGCAAGAGCACAACAATTGTCACGTCCCTCTCGCGTGGA  3903
Qy  1021  TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCGTGT  1080
Db  3904  TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGATCCGTGT  3963
Qy  1081  CCTGTCTTTTAAAACTGCACATTCCCGTGTGAAAGCATTTGGAGACAAGAAGTACAG  1140
Db  3964  CCTGTCTTTTAAAACTGCACATTCCCGTGTGAAAGCATTTGGAGACAAGAAGTACAG  4023
Qy  1141  ATACCTTTTCAAGCAAGTGGCAAGTCTCAAATTGTGACCAGGATTTTGTGACCAAGGATTTTGTGA  1200
Db  4024  ATACCTTTTCAAGCAAGTGGCAAGTCTCAAATTGTGACCAGGATTTTGTGACCAAGGATTTTGTGA  4083
Qy  1201  CCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG  1260
```

```
Qy      1  CGACTTTCCAGCAGTTCAGAAGCAGAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3025  CGACTTTCCAGCAGTTCAGAAGCAGAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA   3084

Qy     61  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3085  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA   3144

Qy    121  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3145  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG   3204

Qy    181  AGCCCAGAATGTCACTCGGCTTCACGAAAACAGGTGAGGAGGTCAATACTGAGTGGGA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3205  AGCCCAGAATGTCACTCGGCTTCACGAAAACAGGTGAGGAGGTCAATACTGAGTGGGA   3264

Qy    241  AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3265  AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT   3324

Qy    301  CCAGGAACTTCAAGAGGCCACGGACTGACTCGAAGCTGCGCCAAGCTGAGGTGAT   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3325  CCAGGAACTTCAAGAGGCCACGGACTGACTCGAAGCTGCGCCAAGCTGAGGTGAT   3384

Qy    361  CAAGGGATCCTGGCAGCCGTGGCGGATCTCCTCATTGACTTCTCCAAGATCACCTCGA   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3385  CAAGGGATCCTGGCAGCCGTGGCGGATCTCCTCATTGACTTCTCCAAGATCACCTCGA   3444

Qy    421  GAAAGTCAAGGCACTTCGAGGAGAAAATTGCGGCTCTGAAAGAAGAACGTGAGCCACGTCAA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3445  GAAAGTCAAGGCACTTCGAGGAGAAAATTGCGGCTCTGAAAGAAGAACGTGAGCCACGTCAA   3504

Qy    481  TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3505  TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC   3564

Qy    541  TCTGGAAGACCTGAACACCAGATGGAGAGCTTCTGCAGGTGCCGTGGACGGACCGAGTCAG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3565  TCTGGAAGACCTGAACACCAGATGGAGAGCTTCTGCAGGTGCCGTGGACGGACCGAGTCAG   3624

Qy    601  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGATCTCAGCACTTCTTCAGCGTC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3625  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGATCTCAGCACTTCTTCAGCGTC   3684

Qy    661  TGTCCAGGGTCCCTGGGAGAGAGCCATCCCAAAATGACAGAGTCTCACCAGTCTTTAGC   720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3685  TGTCCAGGGTCCCTGGGAGAGAGCCATCCCAAAATGACAGAGTCTCACCAGTCTTTAGC   3744

Qy    721  CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTCACCAGTCTTTAGC   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3745  CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTCACCAGTCTTTAGC   3804

Qy    781  TGACCTGAATAAGTCAGATTCTCAGCTTCTTCAGCTCTTGAGCCTGTCAGTCGATGCTGCA   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3805  TGACCTGAATAAGTCAGATTCTCAGCTTCTTCAGCTCTTGAGCCTGTCAGTCGATGCTGCA   3864

Qy    841  GAAGGCCCTTTGCTTGGGATCATCATTTGAGCCTGGCATGCCTTGGACCAGCA   900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3865  GAAGGCCCTTTGCTTGGGATCATCATTTGAGCCTGGCATGCCTTGGACCAGCA   3924

Qy    901  CAACCTCAAGCAGCAAAATGACCAGCCATGGATATCCTGCAGATTATTAATTGTTGACCAC   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3925  CAACCTCAAGCAGCAAAATGACCAGCCATGGATATCCTGCAGATTATTAATTGTTGACCAC   3984

Qy    961  TATTTATGACCGCCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGA   1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3985  TATTTATGACCGCCTGGAGCAAGAGCACAAACAATTTGGTCAACGTCCCTCTCTGCGTGGA   4044

Qy   1021  TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACAGGAGGATCCGTGT   1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4045  TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACAGGAGGATCCGTGT   4104

Qy   1081  CCTGTCTTTTAAAACTGGCATCATTCCCGTGTGTAAAGCACATTTGGAAGACAAGTACAG   1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4105  CCTGTCTTTTAAAACTGGCATCATTCCCGTGTGTAAAGCACATTTGGAAGACAAGTACAG   1140

Qy   1141  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT   1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4047  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT   4106

Qy   1201  CCTTCGCATGCATGATTCTATCCAAATTCCAAGACAGTGGGTGAGTGAAGTTGCATCCTTTGGGGG   1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4107  CCTTCGCATGCATGATTCTATCCAAATTCCAAGACAGTGGGTGAGTGAAGTTGCATCCTTTGGGGG   4166

Qy   1261  CAGTAACATTGAGCCAAGTGTCCCGGAGCTGCTTCCAATTTG   1301
           |||||||||||||||||||||||||||||||||||||||||
Db   4167  CAGTAACATTGAGCCAAGTGTCCCGGAGCTGCTTCCAATTTG   4207
```

RESULT 12
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match   100.0%;   Score 1301;   DB 10;   Length 4966;
Best Local Similarity 100.0%;   Pred. No.0;
Matches 1301;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   481 TGACCTTGCTGCCAGCTTACCACTTGGGCATTCAGCTATAACCTCAGCAC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3529 TGACCTTGCTGCCAGCTTACCACTTGGGCATTCAGCTATAACCTCAGCAC 3588

Qy   541 TCTGGAAGACCTGACACCAGATGGAAGCTTCTGCAGGTGGCCGTGGACCGAGTCAG 600
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   3589 TCTGGAAGACCTGACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGACCGAGTCAG 3648

Qy   601 GCAGCTGCATGAAGCCACCAGGGACTTGGTCCAGCATCCAGCACTTTCTTCCACGTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   3649 GCAGCTGCATGAAGCCACCAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 3708

Qy   661 TGTCCAGGTGCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   3709 TGTCCAGGTGCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3768

Qy   721 CGAGACTCAAACAACTGCTGGACCATCCCAAAATGACACAGAGCTCTACCAGTCTTAGC 780
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   3769 CGAGACTCAAACAACTGCTGGACCATCCCAAAATGACACAGAGCTCTACCAGTCTTTAGC 3828

Qy   781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   3829 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3888

Qy   841 GAAGGGCCCTTGCTTGGATCTCTTGAGCCTGCAGCTGCATGTGATGCCTTGGACCAGCA 900
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   3889 GAAGGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGTGCATGTGATGCCTTGGACCAGCA 3948

Qy   901 CAACCTCAAGCAAAATGACCAGCCCATGGATAATCCGCAGATTATTAATTGTTTGACCAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   3949 CAACCTCAAGCAAAATGACCAGCCCATGGATAATCCGCAGATTATTAATTGTTTGACCAC 4008

Qy   961 TATTTATGACCCCTGGAACTGGCTGCTGAATGTTTATGATACGGGACGAACGAGGAGGATCCGTGT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4009 TATTTATGACCCCTGGAACTGGCTGCTGAATGTTTATGATACGGGACGAACGAGGAGGATCCGTGT 4068

Qy   1021 TATGTGTCTGAACTGGCTGCTGAAGTGTTATGATACGGGACGAACGAGGAGGATCCGTGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4069 TATGTGTCTGAACTGGCTGCTGAAGTGTTATGATACGGGACGAACGAGGAGGATCCGTGT 4128

Qy   1081 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4129 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 4188

Qy   1141 ATACCTTTCAAGCAAGTGGCAAGTTCAAATTCAAATTCTGACAGAGCA 1200
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4189 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGATTTGTGACCAGGCTGGGCCT 4248

Qy   1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTGGGGTGAAGTGCATCCTTTGGGGG 1260
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4249 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTGGGGTGAAGTGCATCCTTTGGGGG 4308

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||
Db   4309 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4349

RESULT 14
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36

Qy   4105 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACGCGGAGTACAG 4164
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATACCTTTCCAGCAGTTCAGAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGGCGGCAGGCTGGGCCT 1200

Qy   4165 ATACCTTTCCAGCAGTTCAGAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGGCGGCAGGCTGGGCCT 4224
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4225 CCTTCTGCATGATTCTGATCTATCCAAATTCCAAGACAGTGGGGTGAAGTTGGGGTGCATCCTTTGGGGG 4284

Qy   4285 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4325
          ||||||||||||||||||||||||||
Db   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301

RESULT 13
US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-34

Query Match   100.0%; Score 1301; DB 10; Length 4990;
Best Local Similarity 100.0%;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 CGACTTTCCAGCAGTTCAGAGCAACGAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3049 CGACTTTCCAGCAGTTCAGAGCAAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3108

Qy   61 AACTAAAGAACCTGTAATCATGATGACTACTTCTTGAGACTGTACGACTTCTGACAGAGCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3109 AACTAAAGAACCTGTAATCATGATGACTACTTCTTGAGACTGTACGACTTCTGACAGAGCA 3168

Qy   121 GCCTTTGGAAGGATCTAGAGAACTAGAGAACTCTACCAGGAGCAGCAGCCCAGAGAGCAGAGAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3169 GCCTTTGGAAGGATCTAGAGAACTAGAGAACTCTACCAGGAGCAGCAGCCCAGAGAGCAGAGAG 3228

Qy   181 AGCCCAGAATGTCACTCGGCTTCTACGGAAAGCAGCAGCTGAGGTCAATACTGAGTGGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3229 AGCCCAGAATGTCACTCGGCTTCTACGGAAAGCAGCAGCTGAGGTCAATACTGAGTGGGA 3288

Qy   241 AAAATTGAACCTGCACTCCGCTGACTGGCTGGCAGGAGAAAATAGATGAGACCCTTGAAAGACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3289 AAAATTGAACCTGCACTCCGCTGACTGGCTGGCAGGAGAAAATAGATGAGACCCTTGAAAGACT 3348

Qy   301 CCAGGAACTTCAAGAGCGCACGGCCACGCCGTGGCTGGACCTGCGCCCAAGCTGCAGGTGAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3349 CCAGGAACTTCAAGAGCGCACGGCCACGCCGTGGCTGGACCTGCGCCCAAGCTGCAGGTGAT 3408

Qy   361 CAAGGGATCCTGGCAGCCGCCGTGGCGATCTCCTCATTGACTTCTCCAAGATCACCTCGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3409 CAAGGGATCCTGGCAGCCGCCGTGGCGATCTCCTCATTGACTTCTCCAAGATCACCTCGA 3468

Qy   421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3469 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAA 3528

```
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match       100.0%;  Score 1301;  DB 10;  Length 5060;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1301;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3119 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3178

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3179 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3238

Qy    121 GCCTTTGGAGGACTAGAGAAACTTCTACCAGGAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3239 GCCTTTGGAGGACTAGAGAAACTTCTACCAGGAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 3298

Qy    181 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGGAGGAGGTCAATACTGAGTGGGA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3299 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGGAGGAGGTCAATACTGAGTGGGA 3358

Qy    241 AAAATTGAACCTGACTCCGCTGGACTGGACGGCAGCAGAGAAAAAATAGATGAGACCCTTGAAAGACT  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3359 AAAATTGAACCTGACTCCGCTGGACTGGACGGCAGCAGAGAAAAAATAGATGAGACCCTTGAAAGACT 3418

Qy    301 CCAGGAACTTCAAGAGCCACGAGGGCCACGAGGGCTGGACCTGGACTGGAGCTGAGGTGAT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3419 CCAGGAACTTCAAGAGCCACGAGGGCCACGAGGGCTGGACCTGGACTGGAGCTGAGGTGAT 3478

Qy    361 CAAGGGATCCTGGCGAGCGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3479 CAAGGGATCCTGGCGAGCGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 3538

Qy    421 GAAAGTCAAGGCACTTCGAGGAGGAGAAATTGCGCCCTGAAAAGAAACGTGAGCCACGTCAA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3539 GAAAGTCAAGGCACTTCGAGGAGGAGAAATTGCGCCCTGAAAAGAAACGTGAGCCACGTCAA 3598

Qy    481 TGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCACGTATAACCTCAGCGAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3599 TGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCACGTATAACCTCAGCGAC 3658

Qy    541 TCTGGAAGAACCTGAACACCAGATGGAAGCTTCTCGCAGGTGGCCGTGGCCGTCTGCGCCGTGCCGAGTCAG  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3659 TCTGGAAGAACCTGAACACCAGATGGAAGCTTCTCGCAGGTGGCCGTGGCCGTCTGCGCCGTGCCGAGTCAG 3718

Qy    601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3719 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3778

Qy    661 TGTCCAGGGGTCCCTGGGAGGAGAGCCATCTGCCAAAACAAAGTGCCCTACTATATCAACCA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3779 TGTCCAGGGGTCCCTGGGAGGAGAGCCATCTGCCAAAACAAAGTGCCCTACTATATCAACCA 3838

Qy    721 CGAGACTCAAACAACTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3839 CGAGACTCAAACAACTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3898

Qy    781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3899 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3958

Qy    841 GAAGGCCCTTTGCTTGCTTGGATCTGTGGGAATCCTGCTTGCTGTCAGCTGCATGTCATGTGATGCCTTGGACCAC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3959 GAAGGCCCTTTGCTTGCTTGGATCTGTGGGAATCCTGCTTGCTGTCAGCTGCATGTCATGTGATGCCTTGGACCAC 4018

Qy    901 CAACCTCAAGCAAAAATGACCACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4019 CAACCTCAAGCAAAAATGACCACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 4078
```

```
Qy    961 TATTTATGACCGCCTGAGCAACAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCGCGTGGA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4079 TATTTATGACCGCCTGAGCAACAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCGCGTGGA 4138

Qy   1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACAAGGGAGGATCCGTGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4139 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACAAGGGAGGATCCGTGT 4198

Qy   1081 CCTGCTCTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4199 CCTGTCTTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG 4258

Qy   1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGCGTGGGCCT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4259 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGCGTGGGCCT 4318

Qy   1201 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4319 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 4378

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGTGCTTCCAATTTG 1301
          ||||||||||||||||||||||||||||||||||||||||
Db   4379 CAGTAACATTGAGCCAAGTGTCCGGAGTGCTTCCAATTTG 4419
```

```
Query Match       100.0%;  Score 1301;  DB 10;  Length 5149;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1301;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3208 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3267

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3268 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3327

Qy    121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3328 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 3387

Qy    181 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTACGAAAGCAGGCTGGAGGGTCAATACTGAGTGGGA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3388 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCTACGAAAGCAGGCTGGAGGGTCAATACTGAGTGGGA 3447

Qy    241 AAAATTGAACCTGACTGGCTGAGAGAAAATAGATGAGATGAGAGACCCTTGAAAGACT  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 AAAATTGAACCTGACTGGCTGAGAGAAAATAGATGAGATGAGAGACCCTTGAAAGACT 3507

Qy    301 CCAGGAACTTCAGAGGCCACGGATGAGCTGGACCTGACCTGACCCAAGTGCGCCAAGTGCGCCAAGTGAGGTGAT  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 CCAGGAACTTCAGAGGCCACGGATGAGCTGGACCTGACCTGACCCAAGTGCGCCAAGTGCGCCAAGTGAGGTGAT 3567
```

```
y    361  CAAGGGATCCTGGCAGCCCGTGGGGCGATCCTCCTCATTGACTCTCTCCAAGATCACCTCGA  420
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
b   3568  CAAGGGATCCTGGCAGCCCGTGGGCGATCCTCCTCATTGACTCTCTCCAAGATCACCTCGA  3627

y    421  GAAAGTCAAGGCACTTCGAGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3628  GAAAGTCAAGGCACTTCGAGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA  3687

y    481  TGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTTCTCACCGTATAACCTCAGCAC  540
          ||||||||||||||||||||||||||||||||||||||   |||||||||||||||||||
b   3688  TGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC  3747

y    541  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3748  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG  3807

y    601  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCCAGCACTTTCTTTCCACGTC  660
          ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
b   3808  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC  3867

y    661  TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTATATCAACCA  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3868  TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTATATCAACCA  3927

y    721  CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3928  CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC  3987

y    781  TGACCTTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   3988  TGACCTTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  4047

y    841  GAAGGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   4048  GAAGGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA  4107

y    901  CAACCTTCAAGCAAAATGACCAGCCCATGAGAATCCTGCGAGATTATTAATTGTTTGACCAC  960
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
b   4108  CAACCTTCAAGCAAAATGACCAGCCCATGAGAATCCTGCAGATTATTAATTGTTTGACCAC  4167

y    961  TATTTATGACCGCCTGAGCAAGAGCACAACAATTGGTCAACGTCCCTCCTCTGCGTGGA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   4168  TATTTATGACCGCCTGAGCAAGAGCACAACAATTGGTCAACGTCCCTCCTCTGCGTGGA  4227

y   1021  TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGGATCCGTGT  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   4228  TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGGATCCGTGT  4287

y   1081  CCTGTCTTTTAAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   4288  CCTGTCTTTTAAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG  4347

y   1141  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCACGCAGGCTGGGCCT  1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   4348  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCACGCAGGCTGGGCCT  4407

y   1201  CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   4408  CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG  4467

y   1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301
          |||||||||||||||||||||||||||||||||||||||||
b   4468  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  4508
```

Search completed: April 5, 2004, 16:33:49
Job time : 322.469 secs

```
              GenCore version 5.1.6
      Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   April  4, 2004, 11:55:32 ; Search time 335.709 Seconds
                                      (without alignments)
                                16463.377 Million cell updates/sec

Title:       US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence:    1 cgacttccagcagttcaga.........ccggagctgcttccaattg 1301

Scoring table: IDENTITY_NUC
            Gapop 10.0 , Gapext 1.0

Searched:   373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:   6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
             1: geneseqn19806:*
             2: geneseqn1990s:*
             3: geneseqn2000s:*
             4: geneseqn2001as:*
             5: geneseqn2001bs:*
             6: geneseqn2002s:*
             7: geneseqn2003as:*
             8: geneseqn2003bs:*
             9: geneseqn2003cs:*
            10: geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID		Description
1	1301	100.0	1821	6	AAD37241		Aad37241 Human dys
2	1301	100.0	2169	6	AAD37232		Aad37232 Human dys
3	1301	100.0	3510	6	AAD37240		Aad37240 Human dys
4	1301	100.0	3531	6	AAD37238		Aad37238 Human dys
5	1301	100.0	3858	6	AAD37237		Aad37237 Human dys
6	1301	100.0	3999	6	AAD37234		Aad37234 Human dys
7	1301	100.0	4182	6	AAD37230		Aad37230 Human dys
8	1301	100.0	4476	6	AAD37259		Aad37259 Adeno-ass
9	1301	100.0	4498	6	AAD37258		Aad37258 Adeno-ass
10	1301	100.0	4825	6	AAD37257		Aad37257 Adeno-ass
11	1301	100.0	4848	6	AAD37263		Aad37263 Adeno-ass
12	1301	100.0	4966	6	AAD37256		Aad37256 Adeno-ass
13	1301	100.0	4990	6	AAD37262		Aad37262 Adeno-ass
14	1301	100.0	5060	6	AAD37264		Aad37264 Adeno-ass
15	1301	100.0	5149	6	AAD37255		Aad37255 Adeno-ass
16	1301	100.0	5339	6	ABK81998		Abk81998 DNA encod
17	1301	100.0	5462	6	ABK81999		Abk81999 DNA encod
18	1301	100.0	5952	5	AAD06794		Aad06794 Human dys
19	1301	100.0	8689	6	ABK82000		Abk82000 DNA encod
20	1301	100.0	11058	6	AAD37229		Aad37229 Human dys
21	1301	100.0	11241	6	ABK82005		Abk82005 cDNA enco
22	1301	100.0	11443	6	ABK82002		Abk82002 DNA encod
23	1301	100.0	12923	1	AAN90338		Aan90338 Sequence
24	1301	100.0	13957	6	ABK81959		Abk81959 cDNA enco
25	1301	100.0	13957	6	ABT10904		Abt10904 Human bre
26	1301	100.0	13957	6	ABN95786		Abn95786 Gene #228
27	1301	100.0	13957	6	ABS69900		Abs69900 Human dys
28	1290	99.2	13977	6	ABS70403		Abs70403 Human bon
29	1135.2	87.3	13815	6	ABK81960		Abk81960 cDNA enco
30	1135.2	87.3	13815	2	ABI99799		Abi99799 Mouse isc
31	1135.2	87.1	19307	2	AAT27558		Aat27558 Shuttle v
32	1133.6	87.1	13815	2	AAV18885		Aav18885 Mus muscu
33	1085.2	83.4	4402	3	AAZ48567		Aaz48567 A rod sho
34	1017.6	78.2	3446	6	AAD37242		Aad37242 Human dys
35	1017.6	78.2	4414	6	AAD37260		Aad37260 Adeno-ass
36	1016.6	78.1	5417	6	ABK81997		Abk81997 DNA encod
37	1016	78.1	1434	6	AAD37243		Aad37243 Human dys
38	886	68.1	4402	3	AAZ48568		Aaz48568 A rod sho
39	750.4	57.7	4075	3	AAZ48569		Aaz48569 A rod sho
40	746.2	57.4	3747	3	AAZ48566		Aaz48566 A rod sho
41	692.6	53.2	3163	3	AAZ48571		Aaz48571 A rod sho
42	582.4	44.8	2005	7	ABT33376		Abt33376 NOVX DNA
43	580.8	44.6	2463	7	ABT33375		Abt33375 NOVX DNA
44	570.4	43.8	10705	7	ABT41896		Abt41896 Toxicity
45	567.2	43.6	11096	6	ABK81962		Abk81962 cDNA enco

ALIGNMENTS

```
RESULT 1
AAD37241
ID   AAD37241 standard; DNA; 1821 BP.
XX
AC   AAD37241;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Human dystrophin rod, hinge and CR domain regions encoding DNA #2.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 52-53; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal of modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
```

```
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin DNA fragment encoding rods R23 and R24, hinge H4 and CR domain
CC   regions
XX
SQ   Sequence 1821 BP; 506 A; 451 C; 447 G; 417 T; 0 U; 0 Other;

Query Match           100.0%;  Score 1301; DB 6; Length 1821;
Best Local Similarity 100.0%;  Pred. No.0;
Matches 1301;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGACATATAGGGCCTTCAAGAGGGAATTGAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGACATATAGGGCCTTCAAGAGGGAATTGAA 162

Qy         61 AACTAAAGAACCTGTAATCATGAGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AACTAAAGAACCTGTAATCATGAGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 222

Qy        121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGAGAGCTGCCTCCTGAGGAGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 GCCTTTGGAAGGACTAGAGAAACTCTACCAGAGAGCTGCCTCCTGAGGAGAG 282

Qy        181 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 AGCCCAGAATGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 342

Qy        241 AAAATTGAACCTGCACTCCGCTGACTGGCGAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 AAAATTGAACCTGCACTCCGCTGACTGGCGAGAGAAAAATAGATGAGACCCTTGAAAGACT 402

Qy        301 CCAGGAACTTCAAGAGGGCCACGATGAGCTGGACCTCAAGCTCGCGCAAGCTGAGGTGAT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 CCAGGAACTTCAAGAGGGCCACGATGAGCTGGACCTCAAGCTCGCGCAAGCTGAGGTGAT 462

Qy        361 CAAGGGATCCTCGGCAGCCCGTGGCGCGATCTCCTCATTGACTCTCCCAAGATCACCTCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 CAAGGGATCCTCGGCAGCCCGTGGCGCGATCTCCTCATTGACTCTCCCAAGATCACCTCGA 522

Qy        421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAGAACGTGAGCCACGTCAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAAGAACGTGAGCCACGTCAA 582

Qy        481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACGGTATAACCTCAGCAC 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACGGTATAACCTCAGCAC 642

Qy        541 TCTGGAAGACCTGAACACACAGAGCTTCTGCAGGTGGAAGCTTCTGCAGGTGGCCGTCGAGGAGACCGAGTCAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 TCTGGAAGACCTGAACACACAGAGCTTCTGCAGGTGGAAGCTTCTGCAGGTGGCCGTCGAGGAGACCGAGTCAG 702

Qy        601 GCAGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCCGCAAACAAAGTGCCCAAAACAAAGTGCCCACTACTATCAACCA 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GCAGCTGCATGAAGCCCACAGGGACTTGGTCCAGCATCCGCAAACAAAGTGCCCAAAACAAAGTGCCCACTACTATCAACCA 762

Qy        661 TGTCCAGGGTCCCTCGGGAGAGAGCCATCCGCAAACAAAGTGCCCACTACTATCAACCA 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 TGTCCAGGGTCCCTCGGGAGAGAGCCATCCGCAAACAAAGTGCCCACTACTATCAACCA 822

Qy        721 CGAGACTCAAACAACACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 CGAGACTCAAACAACACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 882

Qy        781 TGACCTGAATAATGTCAGATTCTCAGCTTCTCAGATCCTGCCATGAAAACTCCGAAGACTGCA 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 TGACCTGAATAATGTCAGATTCTCAGCTTCTCAGATCCTGCCATGAAAACTCCGAAGACTGCA 942

Qy        841 GAAGGCCCTTGCTGGGATCTGTGAGCCTGCAGTGCTGCATGTGATGCCTTGGACCAGCA 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 GAAGGCCCTTGCTGGGATCTGTGAGCCTGCAGTGCTGCATGTGATGCCTTGGACCAGCA 1002

Qy        901 CAACCTCAAGCAAAATGACCAGCCCATGAGATCCTGCAGATTATTAATTGTTTGACCAC 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 CAACCTCAAGCAAAATGACCAGCCCATGAGATCCTGCAGATTATTAATTGTTTGACCAC 1062

Qy        961 TATTTATGACCGGCTGGAGCAGCAAGAGCACAACAATTGGTCAACGTCCTCTGCCTGGA 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 TATTTATGACCGGCTGGAGCAGCAAGAGCACAACAATTGGTCAACGTCCTCTGCCTGGA 1122

Qy       1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACAGGGAGGATCCGTGT 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACAGGGAGGATCCGTGT 1182

Qy       1081 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1242

Qy       1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1302

Qy       1201 CCTTCTGCATGATTCTATCCAAATTCCAAGCAACAGTTGGGTGGAAGTTGCATCCTTTTGGGGG 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 CCTTCTGCATGATTCTATCCAAATTCCAAGCAACAGTTGGGTGGAAGTTGCATCCTTTTGGGGG 1362

Qy       1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
              |||||||||||||||||||||||||||||||||||||||||
Db       1363 CAGTAACATTGAGCCCAAGTGTCCGGAGCCTGCTTCCAATTTG 1403
```

```
RESULT 2
AAD37232
ID   AAD37232 standard; DNA; 2169 BP.
XX
AC   AAD37232;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin rod, hinge and CR domain regions encoding DNA #1.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 45-46; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
```

```
:C  dystrophin DNA fragment encoding rods R22, R23 and R24, hinge H4 and CR
:C  domain regions
:X
:Q  Sequence 2169 BP; 623 A; 529 C; 524 G; 493 T; 0 U; 0 Other;

Query Match    100.0%;  Score 1301;  DB 6;  Length 2169;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1301;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1  CGACTTTCCAGCAGTTCAGAGACGAGAAGCAGAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA    60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   451  CGACTTTCCAGCAGTTCAGAGACGAGAAGCAGAGCAGATGTACATAGGGCCTTCAAGAGGGAATTGAA   510

Qy    61  AACTAAAGAACCTGTAATCATGAGACTCCTTGAGACTGTACGATATTTCTGACAGAGCA   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   511  AACTAAAGAACCTGTAATCATGAGACTCCTTGAGACTGTACGATATTTCTGACAGAGCA   570

Qy   121  GCCTTTGGAAGACTAGAGAAACTCTACCAGGAGCGCCTCCTGAGGAGAG   180
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   571  GCCTTTGGAAGACTAGAGAAACTCTACCAGGAGCGCCTCCTGAGGAGAG   630

Qy   181  AGCCCAGAATGTCACTCGGCTTCTACCGAAAGCAGCTGAGGAGTCAATACTGAGTGGGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   631  AGCCCAGAATGTCACTCGGCTTCTACCGAAAGCAGCTGAGGAGTCAATACTGAGTGGGA   690

Qy   241  AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   691  AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACT   750

Qy   301  CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   751  CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT   810

Qy   361  CAAGGGATCCTGCGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAGATCACCTCGA   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   811  CAAGGGATCCTGCGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAGATCACCTCGA   870

Qy   421  GAAAGTCAAGGCACTTCGAGGAGAAATTGCCGCCTCTGAAGAGAACGTGAGCCACGTCAA   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   871  GAAAGTCAAGGCACTTCGAGGAGAAATTGCCGCCTCTGAAGAGAACGTGAGCCACGTCAA   930

Qy   481  TGACCTTGCTCGCCAGCTTACCACTTTGGGCCATTCAGCTCTCACCGTATAACCTTCAGCAC   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   931  TGACCTTGCTCGCCAGCTTACCACTTTGGGCCATTCAGCTCTCACCGTATAACCTTCAGCAC   990

Qy   541  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   991  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAG  1050

Qy   601  GCAGCTGCATGAGCCCCAGACGTTTGGTCCAGCATCCAGCACTTTCTTCTTCCACGTC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1051  GCAGCTGCATGAGCCCCAGACGTTTGGTCCAGCATCCAGCACTTTCTTCTTCCACGTC  1110

Qy   661  TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1111  TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA  1170

Qy   721  CGAGACTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1171  CGAGACTCAAACAACTTGCTGGACCATCCCAAAATGACAGAGCTTCTACCAGTCTTTAGC  1230

Qy   781  TGACTGGAATAATGTCAGATTCTCAGCTTTATAGGACTGCCATGAAACTCCGAAGACTGCA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1231  TGACTGGAATAATGTCAGATTCTCAGCTTTATAGGACTGCCATGAAACTCCGAAGACTGCA  1290

Qy   841  GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGTCTGCATGTGATGCCTTGGACCAGCA   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1291  GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGTCTGCATGTGATGCCTTGGACCAGCA  1350

Qy   901  CAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1351  CAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC  1410

Qy   961  TATTTATGACCGCCTGGAGCAAGAGCAACAATTGGTCAACGTCCCTCTCTGCGTGGA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1411  TATTTATGACCGCCTGGAGCAAGAGCAACAATTGGTCAACGTCCCTCTCTGCGTGGA  1470

Qy  1021  TATGTGTCTGAACTGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGGATCCGTGT  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1471  TATGTGTCTGAACTGCTGCTGAATGTTATGATACGGGACGAACAGGGAGGGATCCGTGT  1530

Qy  1081  CCTGTCTTTAAAACTGGCATCATTTCCCGTGTAAAGCACATTTGGAGACAAGTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1531  CCTGTCTTTAAAACTGGCATCATTTCCCGTGTAAAGCACATTTGGAGACAAGTACAG  1590

Qy  1141  ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1591  ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT  1650

Qy  1201  CCTTCTGCATGATTCATCCAAATTCCAAGACACAGTTGGGTGAAGTTGCATCCTTTGGGGG  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1651  CCTTCTGCATGATTCATCCAAATTCCAAGACACAGTTGGGTGAAGTTGCATCCTTTGGGGG  1710

Qy  1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301
          |||||||||||||||||||||||||||||||||||||||||
Db  1711  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1751
```

```
RESULT 3
AAD37240
ID  AAD37240 standard; DNA; 3510 BP.
XX
AC  AAD37240;
XX
DT  21-AUG-2002  (first entry)
XX
DE  Human dystrophin minigene delta3510.
XX
KW  Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW  adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW  Becker muscular dystrophy; ds.
XX
OS  Homo sapiens.
XX
PN  WO200183695-A2.
XX
PD  08-NOV-2001.
XX
PF  27-APR-2001; 2001WO-US013677.
XX
PR  28-APR-2000; 2000US-0200777P.
XX
PA  (XIAO/) XIAO X.
XX
PI  Xiao X;
XX
DR  WPI; 2002-049342/06.
XX
PT  New dystrophin minigene for treating Duchenne or Becker muscular
PT  dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT  rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT  gene.
XX
PS  Example 1; Page 51-52; 71pp; English.
XX
CC
CC  The present invention relates to an isolated nucleotide sequence encoding
CC  a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC  terminal domains, rod repeats, utrophin or a spectrin gene. H1 and H4
CC  domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC  invention also relates to a recombinant adeno-associated virus (AAV)
CC  comprising dystrophin minigene operably linked to an expression control
CC  element. The dystrophin minigene in operable linkage with an expression
CC  control element, in a recombinant adeno-associated virus or retrovirus is
CC  useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC  dystrophy (BMD) in a mammalian subject. The present sequence is human
CC  dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
```

```
CC   hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CC   CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1301;  DB 6;  Length 3510;
Best Local Similarity   100.0%;  Pred. No.0;
  Matches 1301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CGACTTTCCAGCAGTTCAGAGAACGACGATGTACATAGGGGCCTTCAAGAGGGAATTGAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1779 CGACTTTCCAGCAGTTCAGAGAACGACGATGTACATAGGGGCCTTCAAGAGGGAATTGAA 1838

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCCTGAGACTGTACGAATATTCTGACAGAGCA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1839 AACTAAAGAACCTGTAATCATGAGTACTCCTGAGACTGTACGAATATTCTGACAGAGCA 1898

Qy    121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGAGAGCTGCCTCCTGAGGAGAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1899 GCCTTTGGAAGGACTAGAGAAACTCTACCAGAGAGCTGCCTCCTGAGGAGAG 1958

Qy    181 AGCCCAGAATGTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1959 AGCCCAGAATGTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2018

Qy    241 AAAATTGAACCTGCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCCTGAAAGACT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2019 AAAATTGAACCTGCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCCCTGAAAGACT 2078

Qy    301 CCAGGAACTTCAAGAGGCCACGATGAGCTGAAGCTTCTGCAGGTGAGCGCCAAGTGAGTGAT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2079 CCAGGAACTTCAAGAGGCCACGATGAGCTGAAGCTTCTGCAGGTGAGCGCCAAGTGAGTGAT 2138

Qy    361 CAAGGGATCCGCAGCCCGTGGGCGAATCTCCTCATTGACTCTCTCCAAGAATCACCTCGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2139 CAAGGGATCCGCAGCCCGTGGGCGAATCTCCTCATTGACTCTCTCCAAGAATCACCTCGA 2198

Qy    421 GAAAGTCAAGGCACTTGAGGAGGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2199 GAAAGTCAAGGCACTTGAGGAGGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 2258

Qy    481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGTCTCCACCGTATAACCTCAGCAC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2259 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGTCTCCACCGTATAACCTCAGCAC 2318

Qy    541 TCTGGAAGACCTGAACACCACAGATGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2319 TCTGGAAGACCTGAACACCACAGATGAAGCTTCTGCAGGTGGCCGTGAGGACCGAGTCAG 2378

Qy    601 GCAGCTGCATGAAGCCACCAGGGACTTTGGTCCAGCATCTCAGCACACTTCTTCCACGTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2379 GCAGCTGCATGAAGCCACCAGGGACTTTGGTCCAGCATCTCAGCACACTTCTTCCACGTC 2438

Qy    661 TGTCCAGGGTCCCTGGGAGGAGAGCCATCTGCGCCAAACAAGTGCCCTACTATATCAACCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2439 TGTCCAGGGTCCCTGGGAGGAGAGCCATCTGCGCCAAACAAGTGCCCTACTATATCAACCA 2498

Qy    721 CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACGAGAGCTCTACCAGTCTTTAGC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2499 CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACGAGAGCTCTACCAGTCTTTAGC 2558

Qy    781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2559 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2618

Qy    841 GAAGGCCCTTGCTTGCTTGGATCTCTGAGCCTGTGAGCTGCAGTGATGCCATGGACCAGCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2619 GAAGGCCCTTGCTTGCTTGGATCTCTGAGCCTGTGAGCTGCAGTGATGCCATGGACCAGCA 2678

Qy    901 CAACTTCAAGCACAAAATGACCAGCCATGGAATATCCTGCAGATTAATTAATTGTTTGACCAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2679 CAACTTCAAGCACAAAATGACCAGCCATGGAATATCCTGCAGATTATTAATTGTTTGACCAC 2738

Qy    961 TATTTATGACCGCCTGGAGCAAGAGACACAACAATTGGTCAACGTCCCTCTCTGGCTGGA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2739 TATTTATGACCGCCTGGAGCAAGAGACACAACAATTGGTCAACGTCCCTCTCTGGCTGGA 2798

Qy   1021 TATGTGTCGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGATCCGTGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2799 TATGTGTCGAACTGGCTGCTGAATGTTATGATACGGGACGAACAGGGAGATCCGTGT 2858

Qy   1081 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   2859 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGCACAAGTACAG 2918

Qy   1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGGCCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2919 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGGCCT 2978

Qy   1201 CCTTCTGCGATTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGCATCCTTTGGGGG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2979 CCTTCTGCGATTCTATCCAAATTCCAAGACAGTTGGTGAAGTTGCATCCTTTGGGGG 3038

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||||||||||||
Db   3039 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 3079
```

```
RESULT 4
AAD37238
ID   AAD37238 standard; DNA; 3531 BP.
XX
AC   AAD37238;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3531.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 50-51; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
```

CC   hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR
CC   domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1301;   DB 6;   Length 3531;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1301;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy    1     CGACTTTCCAGCAGTTCAGAGAGCAGACGATGACATAGGGGCTTCAAGAGGGAATTGAA    60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1800  CGACTTTCCAGCAGTTCAGAGAGCAGACGATGACATAGGGGCTTCAAGAGGGAATTGAA    1859

Qy    61    AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA    120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1860  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACACAGCA    1919

Qy    121   GCCTTTGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCGTGCCTCCTGAGGAGAG    180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1920  GCCTTTGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCGTGCCTCCTGAGGAGAG    1979

Qy    181   AGCCCAGAGAATGTCACTCGGCTTCTACGGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA    240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1980  AGCCCAGAGAATGTCACTCGGCTTCTACGGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA    2039

Qy    241   AAAATTGAACCTGCACTCCGCTGACTGGCTGCAGAGAAAATAGATGAGACCCTTGAAAGACT    300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2040  AAAATTGAACCTGCACTCCGCTGACTGGCTGCAGAGAAAATAGATGAGACCCTTGAAAGACT    2099

Qy    301   CCAGGAACTTCAAGAGGCCACCGATGAGCTGCAAGCTCAAGCTGCGCCAAGTGAGGTGAT    360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2100  CCAGGAACTTCAAGAGGCCACCGATGAGCTGCAAGCTCAAGCTGCGCCAAGTGAGGTGAT    2159

Qy    361   CAAGGGATCCTGCAGCCCGTGGCAGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGA    420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2160  CAAGGGATCCTGCAGCCCGTGGCAGATCTCCTCAATTGACTCTCTCCAAGATCACCTCGA    2219

Qy    421   GAAAGTCAAGGCACTTCAGGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA    480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2220  GAAAGTCAAGGCACTTCAGGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA    2279

Qy    481   TGACCTTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC    540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2280  TGACCTTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC    2339

Qy    541   TCTTGGAAGACCTGAACACCAGATGGAAGCTTCTCGAGGTGGCCGTCGAGGTCGAGTCAG    600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2340  TCTTGGAAGACCTGAACACCAGATGGAAGCTTCTCGAGGTGGCCGTCGAGGTCGAGTCAG    2399

Qy    601   GCAGCTGCATGAGAGCCCCACAGGGACTTGGTCCAGCATCCAGCACTTCTTTCCCACGTC    660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2400  GCAGCTGCATGAGAGCCCCACAGGGACTTGGTCCAGCATCCAGCACTTCTTTCCCACGTC    2459

Qy    661   TGTCCAGGGTCCTGGGAGAGAGCCATCCGCCAAACAAAGTGCCCTACTATATCAACCA    720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2460  TGTCCAGGGTCCTGGGAGAGAGCCATCCGCCAAACAAAGTGCCCTACTATATCAACCA    2519

Qy    721   CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTTACCAGTCTTTAGC    780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2520  CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTTACCAGTCTTTAGC    2579

Qy    781   TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA    840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2580  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA    2639

Qy    841   GAAGGCCCTTTGCTTGGATCTCTTGATCTCTGGAGCTGTCAGTGCATGTGATGCTGCCAGCA    900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2640  GAAGGCCCTTTGCTTGGATCTCTTGATCTCTGGAGCTGTCAGTGCATGTGATGCTGCCAGCA    2699

Qy    901   CAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATCCTGAGATTATTAATTGTTTGACCAC    960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2700  CAACCTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATCCTGAGATTATTAATTGTTTGACCAC    2759

Qy    961   TATTTATGACCGCCTGGAGCAAGAGCAACAATTGGTCAAGCTGCCTCCTGCGTGGA    1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2760  TATTTATGACCGCCTGGAGCAAGAGCAACAATTGGTCAAGCTGCCTCCTGCGTGGA    2819

Qy    1021  TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGGAGGATCCGTGT    1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2820  TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGGAGGATCCGTGT    2879

Qy    1081  CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACACAAGTACAG    1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2880  CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACACAAGTACAG    2939

Qy    1141  ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGCCT    1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2940  ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAGCGCAGGCTGGGGCCT    2999

Qy    1201  CCTTCTGCATGATTCTATCCAAATTCCAAGACACAGTGGGTGGGTGAAGTTGCATCCTTTGGGGG    1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3000  CCTTCTGCATGATTCTATCCAAATTCCAAGACACAGTGGGTGGGTGAAGTTGCATCCTTTGGGGG    3059

Qy    1261  CAGTAACATTGAGCCAAGTGTCCGGAGCGTGCTTCCAATTTG    1301
            ||||||||||||||||||||||||||||||||||||||||||
Db    3060  CAGTAACATTGAGCCAAGTGTCCGGAGCGTGCTTCCAATTTG    3100
```

CC   hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 U; 0 Other;

Query Match   100.0%; Score 1301; DB 6; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2127 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2186

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2187 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2246

Qy    121 GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCTGCCAGAGAGCTGCCTCCTGGAGGAG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2247 GCCTTGGAAGGACTAGAGAAACTCTACCAGGAGCTGCCAGAGAGCTGCCTCCTGGAGGAG 2306

Qy    181 AGCCAGAATGTCACTCGGCTTCTCACGAAAGCAGGCAGGAGGTCAATACTGAGTGGGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2307 AGCCCAGAATGTCACTCGGCTTCTCACGAAAGCAGGCAGGAGGTCAATACTGAGTGGGGA 2366

Qy    241 AAAATTGAACCTGACACTCCCGCTGACTGGCAGAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2367 AAAATTGAACCTGACACTCCGCTGACTGGCAGAGAGAAAAATAGATGAGACCCTTGAAAGACT 2426

Qy    301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCGCAAGCTGAGGTGAT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2427 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCGCAAGCTGAGGTGAT 2486

Qy    361 CAAGGGATCCTGCGACGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2487 CAAGGGATCCTGCGACGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2546

Qy    421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCCACGTCAA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2547 GAAAGTCAAGGCACTTCGAGGAGAAATTGGCGCCTCGAAAGAGAACGTGAGCCACGTCAA 2606

Qy    481 TGACTTGCTCGCGCAGTTACCACTTGGGCATTCGCCGTATAACCTCAGCAC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2607 TGACTTGCTCGCGCAGTTACCACTTGGGCATTCGCCGTATAACCTCAGCAC 2666

Qy    541 TCTGGAAGACCTGAACACACCAGGAAGCTTCTGCAGTGGCCGTGGCCGTCGAGTCAG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2667 TCTGGAAGACCTGAACACACCAGGAAGCTTCTGCAGTGGCCGTGGCCGTCGAGTCAG 2726

Qy    601 GCAGCTGCATGAAGCCCACAGGACTTTGGTCCAGCCATCTCAGCACTTCTTTCCACGTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2727 GCAGCTGCATGAAGCCCACAGGACTTTGGTCCAGCCATCTCAGCACTTCTTTCCACGTC 2786

Qy    661 TGTCCAGGGTCCCTGGGAGAGGAGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCA 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2787 TGTCCAGGGTCCCTGGGAGAGGAGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCA 2846

Qy    721 CGAGACTCAAACAACTGCTGGGACCACCAGAGGTCTTAGC 780
          ||||||||||||||||||||||||||||||||||||||||
Db   2847 CGAGACTCAAACAACTGCTGGGACCACCAGAGGTCTTAGC 2906

Qy    781 TGACCTGAATAATGTCAGATTCCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2907 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAGACTGCA 2966

Qy    841 GAAGGCCCTTTGCTTGGAATCTTGGACCTGCATGTCAGCTGCTGATGCCATGTGATGCCCAGCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2967 GAAGGCCCTTTGCTTGGAATCTTGGACCTGCATGTCAGCTGCTGATGCCATGTGATGCCCAGCA 3026

Qy    901 CAACCTCAAGCAAAATGACCAGCCCATGAATATCCTGCAGATTATTAATTGTTTGACCAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3027 CAACCTCAAGCAAAATGACCAGCCCATGAATATCCTGCAGATTATTAATTGTTTGACCAC 3086

Qy    961 TATTTATGACCGCCTGAGCAAGCAAGAGCACAACAATTTGGTCAAGCAATTTGTCAAGTCCCTCTCTGCGTGGA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3087 TATTTATGACCGCCTGAGCAAGCAAGAGCACAACAATTTGGTCAAGCAATTTGTCAAGTCCCTCTCTGCGTGGA 3146

Qy   1021 TATGTGTCTGAACTGGCTGCTGGAATGTTTATGATACGGAGCAACAGGGAGGATCCGTGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3147 TATGTGTCTGAACTGGCTGCTGGAATGTTTATGATACGGAGCAACAGGGAGGATCCGTGT 3206

Qy   1081 CCTGTCTTTAAAACTGGCACATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3207 CCTGTCTTTAAAACTGGCACATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3266

Qy   1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGCAGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3267 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGCAGATTTTGTGACCAGCGCAGGCTGGGCCCT 3326

Qy   1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3327 CCTTCTGCATGATTCTATCCAAATTCCAAGACACAGTTGGGTGAAGTTGCATCCTTTGGGGG 3386

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||||||||||||
Db   3387 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 3427
```

RESULT 6
AAD37234
ID   AAD37234 standard; DNA; 3999 BP.
XX
AC   AAD37234;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 46-47; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,

```
:C   hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4
:C   and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
:Q   Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 U; 0 Other;

Query Match      100.0%;  Score 1301;  DB 6;  Length 3999;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1301;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

 y     1  CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2268  CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA  2327

 y    61  AACTAAAGAACCTGTAATCATGAGTACTGTCTTGAGACTGTACAGAATATTCTGACAGAGCA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2328  AACTAAAGAACCTGTAATCATGAGTACTGTCTTGAGACTGTACGAATATTTCTGACAGAGCA  2387

 y   121  GCCTTTGGAAGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2388  GCCTTTGGAAGACTAGAGAAACTCTACCAGGAGCCCAGAGAGTGCCTCCTGAGGAGAG  2447

 y   181  AGCCCAGAGAATGTCACTCGGCTTCTACAGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2448  AGCCCAGAGAATGTCACTCGGCTTCTACAGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA  2507

 y   241  AAAATTGAACCTGCACTCCGTCGACTGGCCAGAGAAAAATAGATGAGACCCTTGAAAGACT  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2508  AAAATTGAACCTGCACTCCGTCGACTGGCCAGAGAAAAATAGATGAGACCCTTGAAAGACT  2567

 y   301  CCAGGAACTTCAAGAGGCCACCGACTGGCTGGACATCTGACCTTCAAGCTCGAGGTGAT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2568  CCAGGAACTTCAAGAGGCCACCGACTGGCTGGACATCTGACCTTCAAGCTCGAGGTGAT  2627

 y   361  CAAGGGATCCTGCAGCCCGTGGCCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2628  CAAGGGATCCTGCAGCCCGTGGCCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA  2687

 y   421  GAAAGTCAAGGCACTTCGAGAGAAATTGCGCCTTCGAGAGAGAACGTGAGCCACGTCAA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2688  GAAAGTCAAGGCACTTCGAGAGAAATTGCGCCTTCGAGAGAGAACGTGAGCCACGTCAA  2747

 y   481  TGACCTTGCTCGCCAGCTACCACTTGGGCATTCAGCTCTCCACCGTATAACCTTCAGCAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2748  TGACCTTGCTCGCCAGCTACCACTTGGGCATTCAGCTCTCCACCGTATAACCTTCAGCAC  2807

 y   541  TCTGGAAGACCTGAACACCAGATGGAAGCTTCGCAGGTGGCCGTCGAGGACCGAGTCAG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2808  TCTGGAAGACCTGAACACCAGATGGAAGCTTCGCAGGTGGCCGTCGAGGACCGAGTCAG  2867

 y   601  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCTTCCACGTC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2868  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCTTCCACGTC  2927

 y   661  TGTCCAGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2928  TGTCCAGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA  2987

 y   721  CGAGACTCAAACAACTGCTGGGACCATCCAAATGACAGAGCTCTACCAGTCTTAGC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  2988  CGAGACTCAAACAACTGCTGGGACCATCCAAATGACAGAGCTCTACCAGTCTTAGC  3047

 y   781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  3048  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  3107

 y   841  GAAGGCCCTTTGCTTGGAATCTCTTGAGCCTGTCAGCTGCATGAATGCCTTGGACCAGCA  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  3108  GAAGGCCCTTTGCTTGGAATCTCTTGAGCCTGTCAGCTGCATGAATGCCTTGGACCAGCA  3167

 y   901  CAACCTCAAGCAAAATGACCACCAGCCATGGATATCCTGCAGATTATTAATTGTTGACCAC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  3168  CAACCTCAAGCAAAATGACCACCAGCCATGGATATCCTGCAGATTATTAATTGTTGACCAC  3227

 y   961  TATTTATGACCGCCTGGAGCAAGAGACCACAACAATTGGTCAACGTCCCTCTCTGCGTGGA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  3228  TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTCTGCGTGGA  3287

 y  1021  TATGTGTCTGAACTGGCTGCTGCTGAATGTTTATGATACGGACGAACAGGGGAGATCCGTGT  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  3288  TATGTGTCTGAACTGGCTGCTGCTGAATGTTTATGATACGGACGAACAGGGGAGATCCGTGT  3347

 y  1081  CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGCAAGTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  3348  CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGCAAGTACAG  3407

 y  1141  ATACCTTTCAAGCAAGCAGGTGGCAAGTTCAACAGGATTTTGACCACGCGCAGGCTGGGCCT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  3408  ATACCTTTCAAGCAAGCAGGTGGCAAGTTCAACAGGATTTTGACCACGCGCAGGCTGGGCCT  3467

 y  1201  CCTTCTCGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 b  3468  CCTTCTCGCATGATTCATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG  3527

 y  1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301
          |||||||||||||||||||||||||||||||||||||||||
 b  3528  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  3568
```

RESULT 7
AAD37230
ID   AAD37230 standard; DNA; 4182 BP.
XX
AC   AAD37230;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Human dystrophin minigene delta4173.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 43-44; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is human
CC   dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,

```
CC   hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC   hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids)
XX
SQ   Sequence 4182 BP; 1309 A; 927 C; 976 T; 0 U; Other;

Query Match             100.0%;  Score 1301; DB 6; Length 4182;
Best Local Similarity   100.0%;  Pred. No.0;
Matches 1301;  Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy        1 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGAGGGAATTGAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2451 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGAGGGAATTGAA 2510

Qy       61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2511 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2570

Qy      121 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2571 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 2630

Qy      181 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2631 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2690

Qy      241 AAAATTGAACCTGACTCCGCTGACTGGCAGCAGAGAAAATAGATGAGACCCTTGAAAGACT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2691 AAAATTGAACCTGACTCCGCTGACTGGCAGCAGAGAAAATAGATGAGACCCTTGAAAGACT 2750

Qy      301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTGGAGGTGAT 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2751 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTGGAGGTGAT 2810

Qy      361 CAAGGGATCTCGGCAGCCCGTGGGCCGCATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2811 CAAGGGATCTCGGCAGCCCGTGGGCCGCATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2870

Qy      421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCCTGAAAGAGAACGTGAGCCACGTCAA 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2871 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCCTGAAAGAGAACGTGAGCCACGTCAA 2930

Qy      481 TGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTTCTCACGTATAACCTCAGCAC 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2931 TGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTTCTCACGTATAACCTCAGCAC 2990

Qy      541 TCTGGAAGACCTGAACACCAGATGAAGTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2991 TCTGGAAGACCTGAACACCAGATGAAGTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 3050

Qy      601 GCAGCTGCATGAAGCCCACAGGGCCACCATCTGCCAAACAAAGTGCCCTACTATATCAACCA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3051 GCAGCTGCATGAAGCCCACAGGGCCACCATCTGCCAAACAAAGTGCCCTACTATATCAACCA 3110

Qy      661 TGTCCAGGGGTCCCTGGGAGAGAGCCATCCGCCAAACAAAGTGCCCTACTATATCAACCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3111 TGTCCAGGGGTCCCTGGGAGAGAGCCATCCGCCAAACAAAGTGCCCTACTATATCAACCA 3170

Qy      721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTTAGC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3171 CGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTTAGC 3230

Qy      781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3231 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3290

Qy      841 GAAGGGCCCTTTGCTTGGATCTTCTTGAGCCTGTCAGCTGCCATGTGATGGACCAGCA 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3291 GAAGGGCCCTTTGCTTGGATCTTCTTGAGCCTGTCAGCTGCCATGTGATGGACCAGCA 3350

Qy      901 CAACCTCAAGCAAAATGACCAGCCCATGGAAGATCCTGCAGATTATTAATTGTTGACCAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3351 CAACCTCAAGCAAAATGACCAGCCCATGGAAGATCCTGCAGATTATTAATTGTTGACCAC 3410
```

```
Qy      961 TATTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCTCGCTGCTGGA 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3411 TATTTATGACCGGCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCTCGCTGCTGGA 3470

Qy     1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACGCAACAGGAGGATCCGTGT 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3471 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGAACGCAACAGGAGGATCCGTGT 3530

Qy     1081 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3531 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3590

Qy     1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3591 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 3650

Qy     1201 CCTTCTGCATGATGATCTATCCAAATTCCAAGACACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3651 CCTTCTGCATGATGATCTATCCAAATTCCAAGACACAGTTGGGTGAAGTTGCATCCTTTGGGGG 3710

Qy     1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
            |||||||||||||||||||||||||||||||||||||||||
Db     3711 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 3751
```

```
RESULT 8
AAD37259
ID   AAD37259 standard; DNA; 4476 BP.
XX
AC   AAD37259;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-3510.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 63-65; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
```

```
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 U; 0 Other;

Query Match      100.0%;  Score 1301;  DB 6;  Length 4476;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1301;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2535 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2594

Qy    61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACAGAATATTTCTGACAGAGCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2595 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACAGAATATTTCTGACAGAGCA 2654

Qy   121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGACCCCAGAGAGCTGCCTCCTGAGGAGAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2655 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGACCCCAGAGAGCTGCCTCCTGAGGAGAG 2714

Qy   181 AGCCCAGAATGTCACTCGGCTTCACGAAAAGCAGAGGCTGAGGAGGTCAATACTGAGTGGGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2715 AGCCCAGAATGTCACTCGGCTTCACGAAAAGCAGAGGCTGAGGAGGTCAATACTGAGTGGGA 2774

Qy   241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2775 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 2834

Qy   301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2835 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT 2894

Qy   361 CAAGGGATCCTGCGACGCCCGTGGCCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2895 CAAGGGATCCTGCGACGCCCGTGGCCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGA 2954

Qy   421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2955 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 3014

Qy   481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACGGTATAACCTCAGCAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3015 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACGGTATAACCTCAGCAC 3074

Qy   541 TCTGAAGACCGAACACCAGAGAAGCTTCTGCAGGTGGCCGTCGAGACCGAGTCAG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3075 TCTGAAGACCGAACACCAGAGAAGCTTCTGCAGGTGGCCGTCGAGACCGAGTCAG 3134

Qy   601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3135 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3194

Qy   661 TGTCCAGGGCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3195 TGTCCAGGGCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3254

Qy   721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3255 CGGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACAGAGCTCTACCAGTCTTTAGC 3314

Qy   781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAAGACTGCA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3315 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAAACTCCGAAGACTGCA 3374

Qy   841 GAAGGCCCTTGCTTGGATCTCTTGAGCCTGTCAGCTGCTGCATGTGATGCCTTCGGACCAGCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3375 GAAGGCCCTTGCTTGGATCTCTTGGATCTCTTGAGCCTGTCAGCTGCTGCATGTGATGCCTTCGGACCAGCA 3434

Qy   901 CAACCTCAAGCAGCAAAATGACCAGCCCATGATAATCCTCAGATTATTAATTGTTTGACCAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3435 CAACCTCAAGCAGCAAAATGACCAGCCCATGATAATCCTCAGATTATTAATTGTTTGACCAC 3494

Qy   961 TATTTATGACCGCCTGGAGCAAGAGACCAAGAGACCACAACAATTTGGTCAACGTCCCCTCCTGCGGTGGA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3495 TATTTATGACCGCCTGGAGCAAGAGACAAGAGACCACAACAATTTGGTCAACGTCCCCTCCTGCGGTGGA 3554

Qy  1021 TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGACGAACAGGAGGGATCCGTGT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3555 TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGACGAACAGGAGGGATCCGTGT 3614

Qy  1081 CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3615 CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3674

Qy  1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGCCT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3675 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGCCT 3734

Qy  1201 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3735 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 3794

Qy  1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
           |||||||||||||||||||||||||||||||||||||||||
Db  3795 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 3835

RESULT 9
AAD37258
ID   AAD37258 standard; DNA; 4498 BP.
XX
AC   AAD37258;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-3531.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 62-63; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expresion
CC   control element, in a recombinant adeno-associated virus or retrovirus is
```

```
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 U; 0 Other;

Query Match    100.0%;  Score 1301;  DB 6;  Length 4498;
Best Local Similarity   100.0%;  Pred. No.0;
Matches 1301;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy        1 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGGAGAAGCGATGTACATAGGGGCTTCAAGAGGGAATTGAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2557 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGGAGAAGCGATGTACATAGGGGCTTCAAGAGGGAATTGAA 2616

Qy       61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2617 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2676

Qy      121 GCCTTTGGAAGGACTAGAGAGAAACTTCTACCAGGAGAGACCCAGAGAGCTGCCTCCTGAGGAGAG 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2677 GCCTTTGGAAGGACTAGAGAGAAACTTCTACCAGGAGAGACCCAGAGAGCTGCCTCCTGAGGAGAG 2736

Qy      181 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2737 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2796

Qy      241 AAAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAAATAGAGTGAAAGACCCCTTGAAAGACT 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2797 AAAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAAATAGAGTGAAAGACCCCTTGAAAGACT 2856

Qy      301 CCAGGAACTTCAAGAGGCCACGAGTGAGCTGGACCTCAAGCTGCGCGCCAAGTGGAGTGAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2857 CCAGGAACTTCAAGAGGCCACGAGTGAGCTGGACCTCAAGCTGCGCGCCAAGTGGAGTGAT 2916

Qy      361 CAAGGGATCCTGGCAGCCCGGTGGGCGATCTCCTCAGAGATCACCTCGA 420
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db     2917 CAAGGGATCCTGGCAGCCCGGTGGGCGATCTCCTCAGAGATCACCTCGA 2976

Qy      421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCGCCACGTCAA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2977 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCGAAAGAGAACGTGAGCGCCACGTCAA 3036

Qy      481 TGACCTTGCTCGCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCTCAGCAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3037 TGACCTTGCTCGCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAACCTCAGCAC 3096

Qy      541 TCTGGAAGACCTGAACACCAGAGTGAAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3097 TCTGGAAGACCTGAACACCAGAGTGAAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 3156

Qy      601 GCAGCTGCATGAAGCCACAGGGGACTTGGTCGGAGCATCCAGCACTTTCTTCCCACGTC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3157 GCAGCTGCATGAAGCCACAGGGGACTTGGTCGGAGCATCCAGCACTTTCTTCCCACGTC 3216

Qy      661 TGTCCAGGGTCCCTGGGAGGAGAGCCATCTCGCCAAACAAGAGTGCCTACTATATCAACCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3217 TGTCCAGGGTCCCTGGGAGGAGAGCCATCTCGCCAAACAAGAGTGCCTACTATATCAACCA 3276

Qy      721 CGAGACTTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGACTCTTTAGC 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3277 CGAGACTTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGACTCTTTAGC 3336

Qy      781 TGACCTGAATAATGTCAGATTCAGATTTCTCAGCTTATAGAGACTCCGAAGAACTGCA 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3337 TGACCTGAATAATGTCAGATTCAGATTTCTCAGCTTATAGAGACTCCGAAGAACTGCA 3396

Qy      841 GAAGGGCCCTTGCTTGGATCTTGAGCCTTGCAGCCTGCGATGCGATGGATGCCTTGACCAGCA 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3397 GAAGGGCCCTTGCTTGGATCTTGAGCCTTGCAGCCTGCGATGCGATGGATGCCTTGACCAGCA 3456

Qy      901 CAACCTCAAGCAGCAAATGACCAGCCCATGAATATCCTGCAGATTATTAATTGTTGACCAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3457 CAACCTCAAGCAGCAAATGACCAGCCCATGAATATCCTGCAGATTATTAATTGTTGACCAC 3516

Qy      961 TATTTATGACCGCCTGGACGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTGCCTGGA 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3517 TATTTATGACCGCCTGGACGAGCAAGAGCACAACAATTGGTCAACGTCCCTCTGCCTGGA 3576

Qy     1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACCGGACAACAGGAGGAGGGGGATCCGTGT 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3577 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACCGGACAACAGGAGGAGGGGGATCCGTGT 3636

Qy     1081 CCTGCTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3637 CCTGCTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 3696

Qy     1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGTCAACAGGATTTTGTGACCAGGCCCT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3697 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGTCAACAGGATTTTGTGACCAGGCCCT 3756

Qy     1201 CCTTCTGCATGATCTCATCCAAATTCCAAGACAGTTCGGTGGTGAAGTTGCATCCTTTGGGGG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3757 CCTTCTGCATGATCTCATCCAAATTCCAAGACAGTTCGGTGGTGAAGTTGCATCCTTTGGGGG 3816

Qy     1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
            |||||||||||||||||||||||||||||||||||||||||
Db     3817 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 3857
```

RESULT 10
AAD37257

ID   AAD37257 standard; DNA; 4825 BP.
XX
AC   AAD37257;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular.
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 61-62; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression

CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1301;   DB 6;   Length 4825;
Best Local Similarity 100.0%;   Pred. No.0;
Matches 1301;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGCCTTCAAGAGGGAATTGAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2884 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2943

Qy   61 AACTAAAGAACCTGTAATCATGAGTACTCCTTGAGACTGTACGTACGTGCTCCTGAGAGCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2944 AACTAAAGAACCTGTAATCATGAGTACTCCTTGAGACTGTACGTACGTGCTCCTGAGAGCA 3003

Qy   121 GCCTTGGAGGACTAGAGAAACTCTACCAGAGAGCTGCCTCCTGAGGAGAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3004 GCCTTTGGAAGGACTAGAGAAACTCTACCAGAGAGCTGCCTCCTGAGGAGAG 3063

Qy   181 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3064 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3123

Qy   241 AAAATTGAACCTGACTCCGCTACTCGGCTGGCAGCAGAAAAATAGATGAGACCCTGAAAGACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3124 AAAATTGAACCTGACTCCGCTACTCGGCTGGCAGCAGAAAAATAGATGAGACCCTGAAAGACT 3183

Qy   301 CCAGGAACTTCAAGAGCCACGGATGAGCTGACCTCAAGCTCGCGCAAGTCGAGGTGAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3184 CCAGGAACTTCAAGAGCCACGGATGAGCTGACCTCAAGCTCGCGCAAGTCGAGGTGAT 3243

Qy   361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3244 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA 3303

Qy   421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3304 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 3363

Qy   481 TGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3364 TGACCTTGCTCGCCAGCTTACCACTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 3423

Qy   541 TCTGGAAGACCTGAACACACTTGCTGGACCAGATGGAAGCTTCTGCAGGACGTGGCCGTCGAGGACCGAGTCAG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3424 TCTGGAAGACCTGAACACACTTGCTGGACCAGATGGAAGCTTCTGCAGGACGTGGCCGTCGAGGACCGAGTCAG 3483

Qy   601 GCAGCTGCATGAGACCCAGAGGGACTTTGGTCAGCATCTCAGCACTTCTTCCACGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3484 GCAGCTGCATGAGACCCAGAGGGACTTTGGTCAGCATCTCAGCACTTCTTCCACGTC 3543

Qy   661 TGTCCAGGGTCCCTGGGAGGAGAGCCATCCGCCAAACAAAGTGCCCTACTATATCAACCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3544 TGTCCAGGGTCCCTGGGAGGAGAGCCATCCGCCAAACAAAGTGCCCTACTATATCAACCA 3603

Qy   721 CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3604 CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3663

Qy   781 TGACCTGAATAATGTCAGATTCTCAGCTTCTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3664 TGACCTGAATAATGTCAGATTCTCAGCTTCTATAGGACTGCCATGAAACTCCGAAGACTGCA 3723

Qy   841 GAAGGCCCTTTGCTTGGATCTTGGAGCTCTTGAGCCTGTCAGTCGTGATGTGCATGTGATGCGTCTTGGACCAGCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3724 GAAGGCCCTTTGCTTGGATCTTGGAGCTCTTGAGCCTGTCAGTCGTGATGTGCATGTGATGCGTCTTGGACCAGCA 3783

Qy   901 CAACCTCAAGCAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATGTTGTTGACCAC 960

Db   3784 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATGTTGTTGACCAC 3843
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   961 TATTTATGACCGCCTGGAGCAAGAGCAACAGCACAACAATTGGTCAACGTCCCTCTCGCGTGGA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3844 TATTTATGACCGCCTGGAGCAAGAGCAACAGCACAACAATTGGTCAACGTCCCTCTCGCGTGGA 3903

Qy   1021 TATGTGTTGAACTGGCTGCTGAAATGTTTATGATACGGGAACGAGGGAGGGAGGATCCGTGT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3904 TATGTGTCTGAACTGGCTGCTGAAATGTTTATGATACGGGAACGAGGGAGGGAGGATCCGTGT 3963

Qy   1081 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGAAGACAAGTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3964 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGAAGACAAGTACAG 4023

Qy   1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4024 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCGCAGGCTGGGCCT 4083

Qy   1201 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGGATGAAGTTGCATCCTTTGGGGG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4084 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGGATGAAGTTGCATCCTTTGGGGG 4143

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||||||||||||
Db   4144 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4184

RESULT 11
AAD37263
ID   AAD37263   standard; DNA; 4848 BP.
XX
AC   AAD37263;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Cytomegalovirus.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 68-70; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)

```
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a
CC   cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ   Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 U; 0 Other;

Query Match        100.0%;  Score 1301;  DB 6;  Length 4848;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1301;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 CGACTTCCAGCAGTCAGAGAAGCAGAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2907 CGACTTCCAGCAGTCAGAGAAGCAGAGACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 2966

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2967 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3026

Qy    121 GCCTTTGGAAGGACTAGAGAAAACTCACCAGGAGCGCCCTCCTGAGGAGAG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3027 GCCTTTGGAAGGACTAGAGAAAACTCACCAGGAGCGCCCTCCTGAGGAGAG 3086

Qy    181 AGCCCAGGAGAATGTCACTCGGCTTCACGAAGCAGCTGAGGAGTCAATACTGAGTGGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3087 AGCCCAGGAGAATGTCACTCGGCTTCACGAAGCAGCTGAGGAGTCAATACTGAGTGGGA 3146

Qy    241 AAAATTGAACCTGCACTCCGCTGACTGGCGACAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3147 AAAATTGAACCTGCACTCCGCTGACTGGCGACAGAGAAAAATAGATGAGACCCTTGAAAGACT 3206

Qy    301 CCAGGAACTTCAAGAGCACCGATGAGCTTGACCTCGCGCCAAGCTGAGGTGAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3207 CCAGGAACTTCAAGAGCACCGATGAGCTTGACCTCGCGCCAAGCTGAGGTGAT 3266

Qy    361 CAAGGGATCCTGGCAGCCCGTGGGCCGATCTCCTCATTGACTCTCTCCAAGATCACCTGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3267 CAAGGGATCCTGGCAGCCCGTGGGCCGATCTCCTCATTGACTCTCTCCAAGATCACCTGA 3326

Qy    421 GAAAGTCAAGGCACTCGAGAGGAGAAATTGCCGCCTCGAAAGAGAACGTGAGCCACGTCAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3327 GAAAGTCAAGGCACTCGAGAGGAGAAATTGCCGCCTCGAAAGAGAACGTGAGCCACGTCAA 3386

Qy    481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3387 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC 3446

Qy    541 TCTGGAAGACCTGAACACCACCAGATGGAAGTTCTGCGAGAGTTCTCGAGGACCCAGTCAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3447 TCTGGAAGACCTGAACACCACCAGATGGAAGTTCTGCGAGAGTTCTCGAGGACCCAGTCAG 3506

Qy    601 GCAGCTGCATGAGAGCCCACAGAGGACTTTGGTCCAGATCTCAGCACTTTCTTTCCACGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3507 GCAGCTGCATGAGAGCCCACAGAGGACTTTGGTCCAGATCTCAGCACTTTCTTTCCACGTC 3566

Qy    661 TGTCCAGGGTCCTGGGAGAGGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3567 TGTCCAGGGTCCTGGGAGAGGCCATCTGCCAAACAAAGTGCCCTACTATATCAACCA 3626

Qy    721 CGAGACTCAAAACTTGCTGGGACCATCCTGGACCACCAAAATGACAGAGCTCTTTAGC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3627 CGAGACTCAAAACAACTTGCTGGGACCATCCTGGACCACCAAAATGACAGAGCTCTTTAGC 3686

Qy    781 TGACCTGAATAATGTCAGATTCAGATTCAGCTTATAGGACTGCCATGAAAACTCCGAGACTGCA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3687 TGACCTGAATAATGTCAGATTCAGATTCAGCTTATAGGACTGCCATGAAAACTCCGAGACTGCA 3746

Qy    841 GAAGGCCCTTTGCTTGGATCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3747 GAAGGCCCTTTGCTTGGATCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA 3806

Qy    901 CAACTTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3807 CAACTTCAAGCAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 3866

Qy    961 TATTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAAGTCCCTCTCGCGTGGA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3867 TATTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAAGTCCCTCTCGCGTGGA 3926

Qy   1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCGTGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3927 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCGTGT 3986

Qy   1081 CCTGTCTTTAAAACTGGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3987 CCTGTCTTTAAAACTGGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG 4046

Qy   1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4047 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 4106

Qy   1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTGGGTGAGTTGCATCCTTTGGGGG 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4107 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTGGGTGAGTTGCATCCTTTGGGGG 4166

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAATTTG 1301
          ||||||||||||||||||||||||||||||||||||||||
Db   4167 CAGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAATTTG 4207
```

```
RESULT 12
AAD37256
ID   AAD37256 standard; DNA; 4966 BP.
XX
AC   AAD37256;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 59-60; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
```

CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigenes, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence
XX
SQ   Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1301;   DB 6;   Length 4966;
Best Local Similarity   100.0%;   Pred. No.0;
Matches 1301;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGTGTACATAGGGCCTTCAGAGGGAATTGAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3025 CGACTTTCCAGCAGTTCAGAAGCAGAAGCAGTGTACATAGGGCCTTCAGAGGGAATTGAA 3084

Qy   61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTCTGACAGAGCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3085 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAATATTCTGACAGAGCA 3144

Qy   121 GCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3145 GCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 3204

Qy   181 AGCCCAGAAGTGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3205 AGCCCAGAAGTGTCACTCGGCTTCACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 3264

Qy   241 AAAATTGAACCTGACCACTCCGCTGACTCCGGCAGCAGAGAAAAATAGAGACCCTTGAAAGACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3265 AAAATTGAACCTGACCACTCCGCTGACTCCGGCAGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 3324

Qy   301 CCAGGAACTTCAAGAGCCACACGATGAGCTGGACCTCAAGCTGCGCGCAAGTGAGGTGAT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3325 CCAGGAACTTCAAGAGCCACACGATGAGCTGGACCTCAAGCTGCGCGCAAGTGAGGTGAT 3384

Qy   361 CAAGGGATCCTGGGCAGCCCGTGGGCGTGGGCGATCTCCTCATTGACTCACTCCTCGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3385 CAAGGGATCCTGGGCAGCCCGTGGGCGTGGGCGATCTCCTCATTGACTCACTCCTCGA 3444

Qy   421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCCTCGAAAGAACGTGAGCCACGTCAA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3445 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCCTCGAAAGAACGTGAGCCACGTCAA 3504

Qy   481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGTCTTCACCGTATAACCTCAGCAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3505 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGTCTTCACCGTATAACCTCAGCAC 3564

Qy   541 TCTGGAAGAACTTCAAGACGATGGACACCAGAAGAGTTCTGCAGGTGGCCGTGCAGGAGTCAG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3565 TCTGGAAGAACTTCAAGACGATGGACACCAGAAGAGTTCTGCAGGTGGCCGTGCAGGAGTCAG 3624

Qy   601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3625 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC 3684

Qy   661 TGTCCAGGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3685 TGTCCAGGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3744

Qy   721 CGAGACTCAAACACTTGCTGGGACCAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3745 CGAGACTCAAACAACTTGCTGGGACCAAATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3804

Qy   781 TGACCTGAATAATGTCAGATTTCAGCTTATAGGACTTGCCATGAAACTCCGAAGACTGCA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3805 TGACCTGAATAATGTCAGATTTCAGCTTATAGGACTTGCCATGAAACTCCGAAGACTGCA 3864

Qy   841 GAAGGCCCTTGCTTGGCTGGATCTCTTGGAGCTGCGTCTGTCGCTGCTCAGCTGAATGTGATGCCTTGGACCAGCA 900
           |||||||||

Db   3865 GAAGGCCCTTGCTTGGCTGGATCTCTTGAGCTGCGTCTGTCAGCTGCTCAGCTGAATGTGATGCCTTGGACCAGCA 3924

Qy   901 CAACCTCAAGCAAAATGACCAGCCCATGAATGTTCCTGAGATTATCCTGATATCCTTGACCAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3925 CAACCTCAAGCAAAATGACCAGCCCATGAATGTTCCTGAGATTATCCTGATATTATTGTTGACCAC 3984

Qy   961 TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCCTCTGCGTGGA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3985 TATTTATGACCGCCTGGAGCAAGAGCACAACAATTGGTCAACGTCCTCCTCTGCGTGGA 4044

Qy   1021 TATGTGTCTGAACTGGCTGCTGATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4045 TATGTGTCTGAACTGGCTGCTGATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 4104

Qy   1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4105 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4164

Qy   1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGAGATTTTGTGACCAGCGGCAGGCTGGGCCT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4165 ATACCTTTTCAAGCAAGTGCAAGTTCAACAGGATTTTGTGACCAGCGGCAGGCTGGGCCT 4224

Qy   1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTGGGGG 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4225 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 4284

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAATTTG 1301
           ||||||||||||||||||||||||||||||||||||||||
Db   4285 CAGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAATTTG 4325

RESULT 13
AAD37262
ID   AAD37262 standard; DNA; 4990 BP.
XX
AC   AAD37262;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Cytomegalovirus.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 67-68; 71pp; English.
XX
CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-

```
CC  terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC  domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC  invention also relates to a recombinant adeno-associated virus (AAV)
CC  comprising dystrophin minigene operably linked to an expression control
CC  element. The dystrophin minigene in operable linkage with an expression
CC  control element, in a recombinant adeno-associated virus or retrovirus is
CC  useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC  dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC  vector plasmid construct containing human dystrophin minigene, a
CC  cytomegalovirus (CMV) promoter and a small polyA signal sequence
XX
SQ  Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 U; 0 Other;

Query Match      100.0%;  Score 1301;  DB 6;  Length 4990;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1301;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy         1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3049 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3108

Qy        61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3109 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3168

Qy       121 GCCTTTGAAGGACTAGAGAAACTTCTACCAGGAGCCCAGAGAGCGTGCCTCCTGAGGAGG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3169 GCCTTTGAAGGACTAGAGAAACTTCTACCAGGAGCCCAGAGAGCGTGCCTCCTGAGGAGG 3228

Qy       181 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGAA 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3229 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGAA 3288

Qy       241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAAATGAGAGACCTTGAAAGACT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3289 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAAATGAGAGACCTTGAAAGACT 3348

Qy       301 CCAGGAACTTCAAGAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGTGAGGTGAT 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3349 CCAGGAACTTCAAGAGAGGCCACGGATGAGCTGACCTCAAGCTGCGCCAAGTGAGGTGAT 3408

Qy       361 CAGGGGATCCTGGCAGCCCGTGGGCCAATCTCTCATTGACTCTCTCCAAGATCACCTCGA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3409 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA 3468

Qy       421 GAAAGTCAAGGCACTTCAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3469 GAAAGTCAAGGCACTTCAGGAGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 3528

Qy       481 TGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3529 TGACCTTGCTCGCCAGTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 3588

Qy       541 TCTGGAAGACCTGAACAACAGAGACTTCAGGTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3589 TCTGGAAGACCTGAACAACAGAGACTTCAGGTGCAGGTGGCCGTCGAGGACCGAGTCAG 3648

Qy       601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3649 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGTC 3708

Qy       661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTATATCAACCA 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3709 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCCAAACAAAGTGCCCTACTATATCAACCA 3768

Qy       721 CGAGACTCAAACAACTTGCTCGGGACCACCTGCCTCCCAAAATCCCAAAATGACAGCTCTTAGC 780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3769 CGAGACTCAAACAACTTGCTCGGGACCACCTGCCTCCCAAAATCCCAAAATGACAGCTCTTAGC 3828

Qy       781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3829 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCCGAAGACTGCA 3888
```

```
Qy       841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCCATGTGATGCTTGGACCAGCA 900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3889 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCCATGTGATGCTTGGACCAGCA 3948

Qy       901 CAACCTCAAGCAAATGACCAGCCCATGAATATCCTGCAGATTATTAATTGTTGACCAC 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3949 CAACCTCAAGCAAATGACCAGCCCATGAATATCCTGCAGATTATTAATTGTTGACCAC 4008

Qy       961 TATTTATGACCGCCTGGAGCAAGAGCAACAACAATTGGTCAACGTCCCTCTCGCGTGGA 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4009 TATTTATGACCGCCTGGAGCAAGAGCAACAACAATTGGTCAACGTCCCTCTCGCGTGGA 4068

Qy      1021 TATGTGTCTGAACTGCTGCTGTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4069 TATGTGTCTGAACTGGCTGCTGTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 4128

Qy      1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4129 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4188

Qy      1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4189 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 4248

Qy      1201 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGGGTGAAGTTGCATCCTTTGGGGG 1260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4249 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTTGGGTGGGTGAAGTTGCATCCTTTGGGGG 4308

Qy      1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
             |||||||||||||||||||||||||||||||||||||||||
Db      4309 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4349
```

```
RESULT 14
AAD37264
ID   AAD37264 standard; DNA; 5060 BP.
XX
AC   AAD37264;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Cytomegalovirus.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX
PS   Example 1; Page 70-71; 71pp; English.
XX
```

CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a
XX small polyA signal sequence
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 U; 0 Other;

```
Query Match        100.0%;  Score 1301;  DB 6;  Length 5060;
Best Local Similarity  100.0%;
Matches 1301;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy      1 CGACTTTCCAGAGTTCAGAAGCAGAACGATGTACATAGGGGCTTCAAGAGGGAATTGAA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3119 CGACTTTCCAGAGTTCAGAAGCAGAACGATGTACATAGGGGCTTCAAGAGGGAATTGAA 3178

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3179 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3238

Qy    121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3239 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 3298

Qy    181 AGCCCAGAATGTCACTCGGCTTCACGAAAACAGCGTCGAGGAGGTCAATACTGAGTGGGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3299 AGCCCAGAATGTCACTCGGCTTCACGAAAACAGCGTCGAGGAGGTCAATACTGAGTGGGA 3358

Qy    241 AAAATTGAACCTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTGA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3359 AAAATTGAACCTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTGA 3418

Qy    301 CCAGGAACTTCAAGAGCTGACGGATGAGCTGCGCCAAGCTCAAGCTCAGCTGAGGTGAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3419 CCAGGAACTTCAAGAGCTGACGGATGAGCTGCGCCAAGCTCAAGCTCAGCTGAGGTGAT 3478

Qy    361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3479 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCTCATTGACTCTCTCCAAGATCACCTCGA 3538

Qy    421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3539 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 3598

Qy    481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3599 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 3658

Qy    541 TCTGGAAGACTTGAACACCAGATGAGAAGCTTCTGCGAGGTGGCCGTCGAGGACCGAGTCAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3659 TCTGGAAGACTTGAACACCAGATGGAAGCTTCTGCGAGGTGGCCGTCGAGGACCGAGTCAG 3718

Qy    601 GCAGCTGCATGAAGCCACAGGGACCCACAGGACTTTGGTCCAGCATCCAGCTCTCAGCGTC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3719 GCAGCTGCATGAAGCCACAGGGACCCACAGGACTTTGGTCCAGCATCCAGCTCTCAGCGTC 3778

Qy    661 TGTCCAGGGTCCCTGGGAGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3779 TGTCCAGGGTCCCTGGGAGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 3838

Qy    721 CGAGACTCAAAACAACTGCTGGGACCATCCCAAAATGAACAGAGTCTACCAGTCTTTAGC 780
            ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   3839 CGAGACTCAAACAACTGCTGGGACCATCCCAAAATGAACAGAGTCTACCAGTCTTTAGC 3898

Qy    781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3899 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3958

Qy    841 GAAGGCCCTTGCTTGGATCTTGAGCCTGCCTCAGCTGCATGTGATGCCTTGGACCAGCA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3959 GAAGGCCCTTGCTTGGATCTTGAGCCTGCCTCAGCTGCATGTGATGCCTTGGACCAGCA 4018

Qy    901 CAACCTCAAGCAAATGACCAGCCGCCATGAATATCCTGCAGATTAATAATTGTTGACCAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4019 CAACCTCAAGCAAATGACCAGCCGCCATGAATATCCTGCAGATTAATAATTGTTGACCAC 4078

Qy    961 TATTTATGACCGCCTGCTGAACTGGCTGCTGAATGTTTATGATACGGGGAC... 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4079 TATTTATGACCGCCTGCTGAACTGGCTGCTGAATGTTTATGATACGGGGAC... 4138

Qy   1021 TATGTGTCTGACTGGCTGTGAATGTTTATGTGTGTAAAGCACATTTGGAAGACAAGTACAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4139 TATGTGTCTGACTGGCTGTGAATGTTTATGTGTGTAAAGCACATTTGGAAGACAAGTACAG 4198

Qy   1081 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4199 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4258

Qy   1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAAATTCAAGAACAGTTGGGTGGAAGTTGCA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4259 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 4318

Qy   1201 CCTTCTGCATGATTCATCAAGATGTCCGGAGCTGCTTCCAATTTG 1301
            ||||||||||||||||||||||||||||||||||||||||||||
Db   4319 CCTTCTGCATGATTCATCAAGATGTCCGGAGCTGCTTCCAATTTG 1301

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
            |||||||||||||||||||||||||||||||||||||||||
Db   4379 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4419
```

RESULT 15
AAD37255
ID   AAD37255 standard; DNA; 5149 BP.
XX
AC   AAD37255;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW   Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW   adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW   Becker muscular dystrophy; ds.
XX
OS   Homo sapiens.
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200183695-A2.
XX
PD   08-NOV-2001.
XX
PF   27-APR-2001; 2001WO-US013677.
XX
PR   28-APR-2000; 2000US-0200777P.
XX
PA   (XIAO/) XIAO X.
XX
PI   Xiao X;
XX
DR   WPI; 2002-049342/06.
XX
PT   New dystrophin minigene for treating Duchenne or Becker muscular
PT   dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT   rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
PT   gene.
XX

```
PS   Example 1; Page 57-59; 71pp; English.

XX

CC   The present invention relates to an isolated nucleotide sequence encoding
CC   a dystrophin minigene. The minigene comprises N-terminal or modified N-
CC   terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC   domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC   invention also relates to a recombinant adeno-associated virus (AAV)
CC   comprising dystrophin minigene operably linked to an expression control
CC   element. The dystrophin minigene in operable linkage with an expression
CC   control element, in a recombinant adeno-associated virus or retrovirus is
CC   useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC   dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC   vector plasmid construct containing human dystrophin minigene, a muscle
CC   creatine kinase (MCK) promoter and a small polyA signal sequence

XX

SQ   Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 U; 0 Other;

Query Match        100.0%;   Score 1301;   DB 6;   Length 5149;
Best Local Similarity   100.0%;   Pred. No.0;
Matches 1301;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy      1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGGGGAATTGAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3208 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGGGGAATTGAA 3267

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3268 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3327

Qy    121 GCCTTTGGAAGGACTAGAGAAACTTACCAGGAGAGCCCAGAGAGTGCCTCCTGAGGAGAG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3328 GCCTTTGGAAGGACTAGAGAAACTTACCAGGAGAGCCCAGAGAGTGCCTCCTGAGGAGAG 3387

Qy    181 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGCAGGCTCAATACTGAGTGGGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3388 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGCAGGCTCAATACTGAGTGGGA 3447

Qy    241 AAAATTGAACCTGCACTCCCGCTGACTGGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 AAAATTGAACCTGCACTCCCGCTGACTGGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 3507

Qy    301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAAGCTGAGGTGAT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGCCAAGCTGAGGTGAT 3567

Qy    361 CAAGGGATCCTGCAGCCGTGTGGGCGATCTCCTCATTGACTCTCCTCCAAGATCACCTCGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 CAAGGGATCCTGCAGCCGTGTGGGCGATCTCCTCATTGACTCTCCTCCAAGATCACCTCGA 3627

Qy    421 GAAAGTCAAGGCACTTCGAGAGAATTGCGCCTGAAAAGAACGTGAGCCACGTCAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3628 GAAAGTCAAGGCACTTCGAGAGAATTGCGCCTGAAAAGAACGTGAGCCACGTCAA 3687

Qy    481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3688 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC 3747

Qy    541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCTCAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3748 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCTCAG 3807

Qy    601 GCAGTCGTGAAGCCCATCTCGCAAGATGACCTTCTTCTTCCACGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3808 GCAGTCGTGAAGCCCACATCTCGCAAGATGACCTTCTTCTTCCACGTC 3867

Qy    661 TGTCCAGGGGTCCCTGGGAGGAGAGCCATCTCGCCCAAACAAGTGCCCTACTATATCAACCA 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3868 TGTCCAGGGGTCCCTGGGAGGAGAGCCATCTCGCCCAAACAAGTGCCCTACTATATCAACCA 3927

Qy    721 CGAGACTCAAAACAACTTGCTGGGACCATCCCAAAATGACACCAAAAATGACAGAGTCTTTAGC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3928 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACACCAAAAATGACAGAGTCTTTAGC 3987

Qy    781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGGACTGCCATGAAACTCCGAAGACTGCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3988 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGGACTGCCATGAAACTCCGAAGACTGCA 4047

Qy    841 GAAGGCCCTTGCTTGCTTGGATCTCTTGGAGCGTCAGCTGCAGCTGTCAGTGATGCCTTGGACCAGCA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4048 GAAGGCCCTTGCTTGCTTGGATCTCTTGGAGCGTCAGCTGCAGCTGTCAGTGATGCCTTGGACCAGCA 4107

Qy    901 CAACCTCAAGCAAATGACCAGCCCATGGATATCCTCAGATTATTAATTGTTTGACCAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4108 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 4167

Qy    961 TATTTATGACCGCCTGGAGCAAGCACAACAATTGGTCAACGTCCTCTCTGCGTGGA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4168 TATTTATGACCGCCTGGAGCAAGCACAACAATTGGTCAACGTCCTCTCTGCGTGGA 4227

Qy   1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGCAGGGAGGATCCGTGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4228 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGCAGGGAGGATCCGTGT 4287

Qy   1081 CCTGTCTTTAAAACTGGACATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4288 CCTGTCTTTAAAACTGGACATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4347

Qy   1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTCTGACAGAGCA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   4348 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTCTGACAGAGCA 4407

Qy   1201 CCTTTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAGTTGAAGTTGCATCCTTTGGGGG 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4408 CCTTTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAGTTGAAGTTGCATCCTTTGGGGG 4467

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||||||||||||
Db   4468 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4508


Search completed: April 4, 2004, 18:20:19
Job time : 338.709 secs
```

```
DM nucleic - nucleic search, using sw model

          Copyright   GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

Run on:        April  4, 2004, 12:04:07 ; Search time 3391.8 Seconds
                              (without alignments)
                     16625.161 Million cell updates/sec

Title:         US-09-845-416-10_COPY_1800_3100
Perfect score: 1301
Sequence:      1 cgacttccagcagttcaga.........ccggagctgcttccaattg 1301

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      3470272 seqs, 2167151995 residues

Total number of hits satisfying chosen parameters:    6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
      1: gb_ba:*
      ...
```

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1301	100.0	5339	6	AX538620	AX538620 Sequence
2	1301	100.0	5462	6	AX538621	AX538621 Sequence
3	1301	100.0	5952	6	AR304538	AR304538 Sequence
4	1301	100.0	5952	6	AX114289	AX114289 Sequence
5	1301	100.0	8689	6	AX538622	AX538622 Sequence
6	1301	100.0	11443	6	AX538624	AX538624 Sequence
7	1301	100.0	12057	6	AX538627	AX538627 Sequence
8	1301	100.0	12446	9	HSDMDR	X14298 Human mRNA
9	1301	100.0	13957	6	AX409637	AX409637 Sequence
10	1301	100.0	13957	6	AX538581	AX538581 Sequence
11	1301	100.0	13957	9	HUMDYS	M18533 Homo sapien
12	1290	99.2	13977	6	AR220819	AR220819 Sequence
13	1154.4	88.7	13887	4	AF070485	AF070485 Canis fam
14	1135.2	87.3	13815	6	AX306153	AX306153 Sequence
15	1135.2	87.3	13815	6	AX538582	AX538582 Sequence
16	1135.2	87.3	13815	10	MUSDYSA	M68859 Mouse dystr
17	1135.2	87.3	19307	6	AR093392	AR093392 Sequence
18	1135.2	87.3	19307	6	AR142592	AR142592 Sequence
c 19	1085.2	83.4	4402	6	E30219	E30219 Shortened d
c 20	1016.6	78.1	5417	6	AX538619	AX538619 Sequence
21	904.8	69.5	13575	5	GGDYS	X13369 Chicken mRN
22	886	68.1	4402	6	E30220	E30220 Shortened d
23	750.4	57.7	4075	6	E30221	E30221 Shortened d
24	746.2	57.4	3747	6	E30218	E30218 Shortened d
25	722.4	55.5	2654	5	FSCDYSTRO	M37645 Torpedo cal
26	692.6	53.2	3163	6	E30223	E30223 Shortened d
27	608.2	46.7	3521	5	AF339031	AF339031 Danio rer
28	584	44.9	2110	9	HUMDMDXX	M92650 Human Duche
29	584	44.9	4658	9	BC028720	BC028720 Homo sapi
30	582.4	44.8	2005	6	AX817318	AX817318 Sequence
31	580.8	44.6	2463	6	AX817316	AX817316 Sequence
32	570.4	43.8	10705	10	RNAJ29967	AJ002967 Rattus no
33	567.2	43.6	3161	10	MMGUTRPH	X83506 M.musculus
34	567.2	43.6	11096	6	AX538584	AX538584 Sequence
35	567.2	43.6	11096	10	MMY12229	Y12229 M.musculus
36	553.8	42.6	6045	6	A63605	A63605 Sequence 7
37	553.8	42.6	6045	6	AR281528	AR281528 Sequence
38	553.8	42.6	6059	6	AX107972	AX107972 Sequence
39	553.8	42.6	10302	6	AX538583	AX538583 Sequence
40	553.8	42.6	10302	9	HSMUPS	X69086 H.sapiens m
41	553.8	42.6	10320	6	A63607	A63607 Sequence 9
42	553.8	42.6	10320	6	AR281529	AR281529 Sequence
43	551	42.4	3499	9	HSU43519	U43519 Human dystr
44	551	42.4	5106	6	AX552248	AX552248 Sequence
45	546.8	42.0	3073	10	AF195788	AF195788 Rattus no

ALIGNMENTS

```
RESULT 1
AX538620
LOCUS       AX538620               5339 bp    DNA    linear   PAT 23-NOV-2002
DEFINITION  Sequence 40 from Patent WO0229056.
ACCESSION   AX538620
VERSION     AX538620.1  GI:25271166
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 40 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
```

Pred. No. is the number of results predicted by chance to have a

FEATURES
    source   Location/Qualifiers
                    1..5339
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
                    /note="Synthetic"

ORIGIN

Query Match   100.0%;   Score 1301;   DB 6;   Length 5339;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1301;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAA   60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1993 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAA 2052

Qy   61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAGCA   120
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2053 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAGCA 2112

Qy   121 GCCTTTGGAAGGACTAGAGAAACTTCACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGGAGAG   180
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2113 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 2172

Qy   181 AGCCCAGAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA   240
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2173 AGCCCAGAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2232

Qy   241 AAAATTGAACCTGCACTCCGCTGACTGCCAGAGAAAAATAGATGACCCTTGAAAGACT   300
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2233 AAAATTGAACCTGCACTCCGCTGACTGCCAGAGAAAAATAGATGACCCTTGAAAGACT 2292

Qy   301 CCAGGAACTTCAAGAGAGGCCACGGATGAGCTGGACTTCAGCGCCAAGCTGAGGTGAT   360
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2293 CCAGGAACTTCAAGAGAGGCCACGGATGAGCTGGACTCGCGCCAAGCTGAGGTGAT 2352

Qy   361 CAAGGGATCCTGCAGCCTGCGTGGGCGATCTCCTCATTGACTCTCCAAGATCACCTCGA   420
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2353 CAAGGGATCCTGCAGCCTGCGTGGGCGATCTCCTCATTGACTCTCCAAGATCACCTCGA 2412

Qy   421 GAAAGTCAAGGCACTTCGAGAGGAGAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAA   480
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2413 GAAAGTCAAGGCACTTCGAGAGGAGAAATTGCGCCTCTGAAAAGAGAACGTGAGCCACGTCAA 2472

Qy   481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCACCGTATAACCTCAGCAC   540
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2473 TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCACCGTATAACCTCAGCAC 2532

Qy   541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGAGGACCGAGTCAG   600
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2533 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGAGGACCGAGTCAG 2592

Qy   601 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTC   660
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2593 GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTCTTCCACGTC 2652

Qy   661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA   720
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2653 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2712

Qy   721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCACCAGTCTTTAGC   780
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2713 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTTCACCAGTCTTTAGC 2772

Qy   781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA   840
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2773 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 2832

Qy   841 GAAGGCCCTTGCTTGGACTCTCTTGAGCCTGTCAGCTGTCAGCCTGTCATGTGATGCCACCAGCA   900
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2833 GAAGGCCCTTGCTTGGACTCTCTTGAGCCTGTCAGCTGTCAGCCTGTCATGTGATGCCACCAGCA 2892

Qy   901 CAACCTCAAGCAAATGACCAGCCCATGGAATCCTGCAGATTATTAATTGTTTGACCAC   960
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2893 CAACCTCAAGCAAATGACCAGCCCATGGAATCCTGCAGATTATTAATTGTTTGACCAC 2952

Qy   961 TATTTATGACCGCCTGGACGAAGAGCAACAACATTTGGTCAAGCTCCCTCTGCGTGGA   1020
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2953 TATTTATGACCGCCTGGACGAAGAGCACACAACAATTGGTCAACGTCCCTCTGCGTGGA 3012

Qy   1021 TATGTGTCTGAACTGGCTGCTGATGTTTATGATACGGGACGAACAGGGGATCCGTGT   1080
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3013 TATGTGTCTGAACTGGCTGCTGATGTTTATGATACGGGACGAACAGGGGATCCGTGT 3072

Qy   1081 CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTGAAGACACATTTGGAAGACAAGTACAG   1140
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3073 CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTGAAAGCACATTTGGAAGACAAGTACAG 3132

Qy   1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCCGAGGCTGGGCCT   1200
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3133 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCCGAGGCTGGGCCT 3192

Qy   1201 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG   1260
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3193 CCTTCTGCATGATTCATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 3252

Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG   1301
             |||||||||||||||||||||||||||||||||||||||||
Db   3253 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 3293

RESULT 2
AX538621
LOCUS   AX538621   5462 bp   DNA   linear   PAT   23-NOV-2002
DEFINITION   Sequence 41 from Patent WO0229056.
ACCESSION   AX538621
VERSION   AX538621.1   GI:25271168
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE   Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 41 11-APR-2002;
             THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
    source   1..5462
                  /organism="synthetic construct"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:32630"
                  /note="Synthetic"
ORIGIN

Query Match   100.0%;   Score 1301;   DB 6;   Length 5462;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1301;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAA   60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2116 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGGAATTGAA 2175

Qy   61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAGCA   120
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2176 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAGAGCA 2235

Qy   121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGGAGAG   180
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2236 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 2295

Qy   181 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA   240
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2296 AGCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 2355

Qy   241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGACCCTTGAAAGACT   300
             ||||||||||||||||||||||||||||||||||||||||||||||||

```
LOCUS           AR304538                    5952 bp    DNA            linear    PAT 12-JUN-2003
DEFINITION      Sequence 1 from patent US 6544786.
ACCESSION       AR304538
VERSION         AR304538.1    GI:31693691
KEYWORDS        .
SOURCE          Unknown.
  ORGANISM      Unknown.
                Unclassified.
REFERENCE       1  (bases 1 to 5952)
  AUTHORS       Xiao,X. and Liu,P.X.
  TITLE         Method and vector for producing and transferring trans-spliced
                peptides
  JOURNAL       Patent: US 6544786-A 1 08-APR-2003;
FEATURES             Location/Qualifiers
     source          1..5952
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match          100.0%;  Score 1301;  DB 6;  Length 5952;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3402 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 3461

Qy       61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTCTTGAGACTCTTGAGACTTCTGACACAGAGCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3462 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTCTTGAGACTTCTGACACAGAGCA 3521

Qy      121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGAGAGCCCAGAGCTGCCTCCTGAGGAGAG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3522 GCCTTTGGAAGGACTAGAGAAACTCTACCAGAGAGCCCAGAGCTGCCTCCTGAGGAGAG 3581

Qy      181 AGCCCAGAATGTCACTCGGCTTCTCACGGCTTCTCACGGCTTCTCTACGAAAGACTGGGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3582 AGCCCAGAATGTCACTCGGCTTCTCACGGCTTCTCTACGAAAGACTGGGA 3641

Qy      241 AAAATTGAACCTGCACTCCGTCACTCGGCAGGAGAAAAATAGATGAGACCCTTGAAAGACT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3642 AAAATTGAACCTGCACTCCGTCGCAGGAGAAAAATAGATGAGACCCTTGAAAGACT 3701

Qy      301 CCAGGAACTTCAAGAGCCACGATGAGCTGGACCTGAGGTGAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3702 CCAGGAACTTCAAGAGCCACGATGAGCTGGACCTGAGGTGAT 3761

Qy      361 CAAGGGATCCTGGCAGCCCGTGGGCAACAGGGAGATCCGTGGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3762 CAAGGGATCCTGGCAGCCCGTGGGCAACAGGGAGATCCGTGGA 3821

Qy      421 GAAAGTCAAGGCACTTGCAGCTTACCTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3822 GAAAGTCAAGGCACTTGCAGCTTACCTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 3881

Qy      481 TGACCTTGCTCGCCAGCTTACCACTTTGGTGTCCAGGTGGGACCGAGGACCGAGGTGAT 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3882 TGACCTTGCTCGCCAGCTTACCACTTTGGTGTCCAGGTGGGACCGAGGACCGAGGTGAT 3941

Qy      541 TCTGGAAGACCTGAACGATGGGAAGCTTGGTCCAGCATCTCAGCACTTTCTTCCCAGTC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3942 TCTGGAAGACCTGAACGATGGGAAGCTTGGTCCAGCATCTCAGCACTTTCTTCCCAGTC 4001

Qy      601 GCAGCTGCATGAAGCCCAGGGACCATCTCGCCAAACAACAAAGTGCCCTACTATATCAACCA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4002 GCAGCTGCATGAAGCCCAGGGACCATCTCGCCAAACAACAAAGTGCCCTACTATATCAACCA 4061

Qy      661 TGTCCAGGGTCCCTGGGAGAGCCATCCTGGGAGAGCCATCCTGGGAGACTCTTTAGC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4062 TGTCCAGGGTCCCTGGGAGAGCCATCCTGGGAGAGCCATCCTGGGACTACTATATCAACCA 4121

Qy      721 CGAGACTCAAACAACTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 3
AR304538

```
Db  3402 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 3461
Qy    61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3462 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 3521

Qy   121 GCCTTGGAAGGACTAGAGAAACTCTACCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3522 GCCTTGGAAGGACTAGAGAAACTCTACCAGAGCCCAGAGAGCTGCCTCCTGAGGAGAG 3581

Qy   181 AGCCCAGAATGTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3582 AGCCCAGAATGTCACTCGGCTTCTCACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA 3641

Qy   241 AAAATTGAACCTGCACTCCGTGACTGGCGAGCAGAGAAAATAGATGAGACCCTTGAAAGACT 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3642 AAAATTGAACCTGCACTCCGTGACTGGCGAGCAGAGAAAATAGATGAGACCCTTGAAAGACT 3701

Qy   301 CCAGGAACTTCAAGAGCCACGGATGAGCTGGACTCCCATTGACTCTCTCCAAGATCACCTGA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3702 CCAGGAACTTCAAGAGCCACGGATGAGCTGGACTCCCATTGACTCTCTCCAAGATCACCTGA 3761

Qy   361 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTGTGTAAAGAACAATTTGAAGACAAGTACAG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3762 CAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTGTGTAAAGAACAATTTGAAGACAAGTACAG 3821

Qy   421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGGCCTTCGAAAAGAGAACTGAGCGACCACGTCAA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3822 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGGCCTTCGAAAAGAGAACGTGAGCGACCACGTCAA 3881

Qy   481 TGAACTTGCTGCCAGCTTACCACTTGGGCATTCAGCTCACGGTATAACCTCAGCAC 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3882 TGACCTTGCTGCCAGCTTACCACTTGGGCATTCAGCTCACCGTATAACCTCAGCAC 3941

Qy   541 TCTGGAAGACTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3942 TCTGGAAGACTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAG 4001

Qy   601 GCAGCTGCATGAACGCCACAGGACTTTGGTCCAGCATCCAGCACTTCTTCAGCGTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4002 GCAGCTGCATGAACGCCCACAGGACTTTGGTCCAGCATCCAGCACTTCTTCAGCGTC 4061

Qy   661 TGTCCAGGGTCCCTGGGAGAGAGCCATCCGCCAAACAAGTGCCCTACTATATCACCA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4062 TGTCCAGGGTCCCTGGGAGAGAGCCATCCGCCAAACAAAGTGCCCTACTATATCACCA 4121

Qy   721 CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4122 CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 4181
```

```
Db  4122 CGAGACTCAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC 4181
Qy   781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4182 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 4241

Qy   841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4242 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCAGCA 4301

Qy   901 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4302 CAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTGTTTGACCAC 4361

Qy   961 TATTTATGACCGGCCTGGAGCAAGAGCACAACAATTTGGTCAACAGCACACAACAGTCCGTGA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4362 TATTTATGACCGGCCTGGAGCAAGAGCACAACAATTTGGTCAACAGCACACAACAGTCCGTGA 4421

Qy  1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACATTTGAAAGACAATTTGACCGT 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4422 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACATTTGAAAGACAATTTGCCGTGT 4481

Qy  1081 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCCATTTGAAAGACAATTTGAAAGAC 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4482 CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCCATTTGAAAGACAATTTGAAAGAC 4541

Qy  1141 ATACCCTTTCAAGCAAGTGGCAAGTOGCAAGTTGGGCCT 1200
         |||||||||||||||||||||||||||||||||||||||
Db  4542 ATACCTTTTCAAGCAAGTGGCAAGTTGGGCCT 4601

Qy  1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTGGGGG 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4602 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTGGGGG 4661

Qy  1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
         |||||||||||||||||||||||||||||||||||||||||
Db  4662 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4702
```

```
Query Match       100.0%;   Score 1301;   DB 6;   Length 5952;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 1301;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy  1 CGACTTTCCGCAGTCAGTTCAGAGAGCAGAACGATGTACATAGGGCCTTCAAGAGGGAATTGAA 60
```

```
Db   4482 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4541
Qy   1141 ATACCTTTTCAAGCAAGTGGCAAGTTCATCCAAATTCAAGACAGGATTTTGTGACCACGGCTCGGCCCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4542 ATACCTTTTCAAGCAAGTGGCAAGTTCATCCAAATTCAAGACAGGATTTTGTGACCACGGCTCGGCCCT 4601
Qy   1201 CCTTCTGCATGATTCTATCCAAGACAGTGTCCGGAGCTGCTTCCATCCTTTGGGGG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4602 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTGTCCGGAGCTGCTTCCATCCTTTGGGGG 4661
Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||||||||||||
Db   4662 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4702
```

RESULT 5
```
AX538622
LOCUS      AX538622               8689 bp   DNA       linear   PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION  AX538622
VERSION    AX538622.1  GI:25271171
KEYWORDS   .
SOURCE     synthetic construct
  ORGANISM synthetic construct
           artificial sequences.
REFERENCE  1
  AUTHORS  Chamberlain,J.S. and Harper,S.Q.
  TITLE    Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL  Patent: WO 0229056-A 42 11-APR-2002;
           THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
  source          1..8689
                  /organism="synthetic construct"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:32630"
                  /note="Synthetic"
ORIGIN
```

```
Query Match      100.0%;  Score 1301;  DB 6;  Length 8689;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1301;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGACGGAATTGAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGACGGAATTGAA 3507
Qy     61 AACTAAAGAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 AACTAAAGAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3567
Qy    121 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGAG 3627
Qy    181 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCACGAAAGCAGGCTGGAGGAGGTCAATACTGAGTGGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3628 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCACGAAAGCAGGCTGGAGGAGGTCAATACTGAGTGGGA 3687
Qy    241 AAAATTGAACCTGACTCGGCTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3688 AAAATTGAACCTGACTCGGCTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 3747
Qy    301 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGAGGTGAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3748 CCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGGAGGTGAT 3807
Qy    361 CAAGGGATCCTGGCGAGCCCGTGGCGATCCTCCTCATTGACTCCTCCAAGATCACCTCGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3808 CAAGGGATCCTGGCGAGCCCGTGGCGATCCTCCTCATTGACTCCTCCAAGATCACCTCGA 3867
Qy    421 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTCACGCCACGTCAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3868 GAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTCACGCCACGTCAA 3927
```

```
Qy    481 TGACCTTGCTGCCAGCTTACCACTTGGGCATTCAGCTCTCCACGTATAACCTCAGCAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3928 TGACCTTGCTGCCAGCTTACCACTTGGGCATTCAGCTCTCCACGTATAACCTCAGCAC 3987
Qy    541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGCTGGACCGAGTCAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3988 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGCTGGACCGAGTCAG 4047
Qy    601 GCAGCTGCATGAGCCCACCAGGGACTTTGGTCCAGCATCCAGCACTTCTTTCCACGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4048 GCAGCTGCATGAGCCCACCAGGGACTTTGGTCCAGCATCCAGCACTTCTTTCCACGTC 4107
Qy    661 TGTCCAGGTGCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4108 TGTCCAGGTGCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 4167
Qy    721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTCCAGTCTTTAGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4168 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTCCAGTCTTTAGC 4227
Qy    781 TGACCTGAATAATGTCAGAATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4228 TGACCTGAATAATGTCAGAATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 4287
Qy    841 GAAGGCCCTTGCTTGGATCTCTGAGCCGTGTCAGCGTGATGCCTTGGACCAGGA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4288 GAAGGCCCTTGCTTGGATCTCTGAGCCGTGTCAGCGTGATGCCTTGGACCAGCA 4347
Qy    901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTCGAGATTATTAATTGTTTGACCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4348 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTCGAGATTATTAATTGTTTGACCAC 4407
Qy    961 TATTTATGACCGCCTGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTGCGTGGA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4408 TATTTATGACCGCCTGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTGCGTGGA 4467
Qy   1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACAACAGGGAGGATCCGTGT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4468 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACAACAGGGAGGATCCGTGT 4527
Qy   1081 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4528 CCTGTCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4587
Qy   1141 ATACCTTTTCAAGCAAGTGGCAAGTTTTGTGACCAGATTTTGTGACCAGGCTGGGCCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4588 ATACCTTTTCAAGCAAGTGGCAAGTTTTGTGACCAGATTTTGTGACACAAGTACAG 4647
Qy   1201 CCTTCTGCATGATTCTATCCAAATTCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4648 CCTTCTGCATGATTCTATCCAAATTCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 4707
Qy   1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||||||||||||
Db   4708 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 4748
```

RESULT 6
```
AX538624
LOCUS      AX538624              11443 bp   DNA       linear   PAT 23-NOV-2002
DEFINITION Sequence 44 from Patent WO0229056.
ACCESSION  AX538624
VERSION    AX538624.1  GI:25271175
KEYWORDS   .
SOURCE     synthetic construct
  ORGANISM synthetic construct
           artificial sequences.
REFERENCE  1
  AUTHORS  Chamberlain,J.S. and Harper,S.Q.
  TITLE    Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL  Patent: WO 0229056-A 44 11-APR-2002;
           THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
```

```
FEATURES       Location/Qualifiers
     source    1..11443
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Synthetic"
ORIGIN
Query Match    100.0%;   Score 1301;   DB 6;   Length 11443;
Best Local Similarity 100.0%;  Pred. No. 0;          Gaps   0;
Matches 1301;   Conservative 0;  Mismatches 0;  Indels 0;  Gaps   0;

Qy       1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6202 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAA 6261

Qy      61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6262 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 6321

Qy     121 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6322 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGAG 6381

Qy     181 AGCCCAGAAGTGTCAATGTCTTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6382 AGCCCAGAAGTGTCAATGTCTTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGA 6441

Qy     241 AAAATGAACCTGCACTCCGCTGACTGGCAGTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6442 AAAATTGAACCTGCACTCCGCTGACTGGCAGTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 6501

Qy     301 CCAGGAACTTCAAGAGAGGCCACGGATGAGTCAAGCTGACCTCAAGCTGGAGGTGAT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6502 CCAGGAACTTCAAGAGAGGCCACGGATGAGTCAAGCTGACCTCAAGCTGGAGGTGAT 6561

Qy     361 CAAGGATCCTGCGCAGCCCGTGGGGCGATCTCATTGACTCTCTCCAAGATCACCTCGA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6562 CAAGGATCCTGCGCAGCCCGTGGGGCGATCTCATTGACTCTCTCCAAGATCACCTCCGA 6621

Qy     421 GAAAGTCAAGGCACTTCGAGAGAAATTGCGCCCTCGAAAGAGAGCGTGAGCCACGTCAA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6622 GAAAGTCAAGGCACTTCGAGAGAAATTGCGCCCTCGAAAGAGAGCGTGAGCCACGTCAA 6681

Qy     481 TGACCTTGCTCGCGCAGCTTACCACTTGGGCATTCAGCTTCCACCGTATAACCTCAGCAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6682 TGACCTTGCTCGCGCAGCTTACCACTTTGGGCATTCAGCTTCCACCGTATAACCTCAGCAC 6741

Qy     541 TCTGGAAGACTTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGCGGTGCCGTGAGTCAG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6742 TCTGGAAGACTTGAACACCAGATGGAAGCTTCTGCAGGTGGCCGTGCGGTGCCGTGAGTCAG 6801

Qy     601 GCAGCTGCGATGAAGCCCACAGGGACTTTGGTCCAGCATCTCCAGCACTTCTTTCTTCCACGTC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6802 GCAGCTGCGATGAAGCCCACAGGGACTTTGGTCCAGCATCTCCAGCACTTCTTTCTTCCACGTC 6861

Qy     661 TGTCCAGGGTCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTATATCAACCA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6862 TGTCCAGGGTCCTGGGAGAGAGCCATCTCGCCAAACAAGTGCCCTACTATATCAACCA 6921

Qy     721 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTACCAGTCTTTAGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6922 CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGTCTACCAGTCTTTAGC 6981

Qy     781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6982 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 7041

Qy     841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCGTCATGTGATGTGCCTGGACCAGCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7042 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCGTCATGTGATGTGCCTGGACCAGCA 7101

Qy     901 CAACCTCAGCAAATGACCAGCCCATGGCATCCTGCAGATTATTAATTGTTTGACCAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7102 CAACCTCAGCAAATGACCAGCCCATGGCATCCTGCAGATTATTAATTGTTTGACCAC 7161

Qy     961 TATTTATGACCGCCTGGGACGAACAAGAGCACAACACATTTGGTCAACGTCCCTCTGCGTGGA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7162 TATTTATGACCGCCTGGGACGAACAAGAGCACAACACATTTGGTCAACGTCCCTCTGCGTGGA 7221

Qy    1021 TATGTGTCTGAACTGGCTGCTGAAGTGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7222 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGT 7281

Qy    1081 CCTGTCTTTTAAAACTGGCATCATTTCCCGTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7282 CCTGTCTTTTAAAACTGGCATCATTTCCCGTGTGTAAAGCACATTTGGAAGACAAGTACAG 7341

Qy    1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7342 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGGCTGGGCCT 7401

Qy    1201 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTGGGGG 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7402 CCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGG 7461

Qy    1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 1301
           |||||||||||||||||||||||||||||||||||||||||
Db    7462 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG 7502
```

```
RESULT 7
AX538627
LOCUS        AX538627              12057 bp    DNA     linear    PAT 23-NOV-2002
DEFINITION   Sequence 47 from Patent WO0229056.
ACCESSION    AX538627
VERSION      AX538627.1  GI:25271181
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    Chamberlain,J.S. and Harper,S.Q.
  TITLE      Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL    Patent: WO 0229056-A 47 11-APR-2002;
             THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES       Location/Qualifiers
     source    1..12057
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Synthetic"
ORIGIN
Query Match    100.0%;   Score 1301;   DB 6;   Length 12057;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1301;   Conservative 0;  Mismatches 0;  Indels 0;  Gaps   0;

Qy       1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8716 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGAGGGAATTGAA 8775

Qy      61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8776 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 8835

Qy     121 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8836 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGGCCCAGAGAGCTGCCTCCTGAGGAGAG 8895

Qy     181 AGCCCAGAAGTGTCAATGTCTTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8896 AGCCCAGAAGTGTCAATGTCTTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGA 8955

Qy     241 AAAATTGAACCTGCACTCCGCTGACTGGCAGTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT 300
```

```
LOCUS       HSDMDR                 12446 bp    mRNA    linear   PRI 12-SEP-1993
DEFINITION  Human mRNA for dystrophin.
ACCESSION   X14298
VERSION     X14298.1  GI:30845
KEYWORDS    Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 12446)
  AUTHORS   Rosenthal,A., Speer,A., Billwitz,H., Cross,G.S., Forrest,S.M. and
            Davies,K.E.
  TITLE     Two human cDNA molecules coding for the Duchenne muscular dystrophy
            (DMD) locus are highly homologous
  JOURNAL   Nucleic Acids Res. 17 (13), 5391 (1989)
  MEDLINE   89345106
   PUBMED   2668885
REFERENCE   2  (bases 1 to 12446)
  AUTHORS   Rosenthal,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
            der DDr, Zentralinstitut fuer Molekularbiologie, Robert-Roessle
            Str.10, 1115 Berlin Buch, DDR
COMMENT     see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
            M18533: cDNA were located at x14298 pos. 496, 1772, 1965, 2449,
            3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES             Location/Qualifiers
     source          1..12446
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X chromosomal, Xp21."
                     /tissue_type="muscle"
                     /dev_stage="fetal and adult."
     CDS             99..11156
                     /note="unnamed protein product; dystrophin (AA 1 - 3685)"
                     /codon_start=1
                     /protein_id="CAA32479.1"
                     /db_xref="GI:30846"
                     /db_xref="GOA:P11532"
                     /db_xref="SWISS-PROT:P11532"
                     /translation="MLWWEEVEDCYEREDVCKKFFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTGLIWNIILHWQVKNVMKNIMAGLQPTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEHFQLHHQMHYSQ
                     QITVSLAQGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFPSQHLEAPEDKSFGSSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTLQAQGEISNDWVKVQDIETHTHEYMLDLTAH
                     QGRVGNIIQLGSKLLIGTKSLSDEETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLKELNDWLTKTEERTRKMEEEPLGPDLEDLKRVQVQHKVLQDLEQEQVRVN
                     SLTHMVVVDESSGDHATAALEEEQLKVLGDRWANICRWTEDRWVLLQDILLKWQRLTE
                     EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLQKLAVLKADLEKKKQSMGKLYSIK
                     QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISQAVTTTQPSLTQTTVM
                     ETVTTVTTREQILVKHAQEELPPPPQKKRQITVDSEIRKRLDVDITELHSWITRSEA
                     VLQSPEFAIFRKEGNFSDLKEKVNAIEREKAEKFRKLQDASRSGQALVEQMVNEGVNA
                     DSIKQASEQLNSRWIEPCQLLSERLNWLEYQNNIIAFYNGLQQLEQMTTTABANWLKIQ
                     PTTPSEPTAIKSQLKICKDEVNRLSGLQPQIERLKIQSIALKEKGGQPMFLDADFVAF
                     TNHFKQVFSDVQAREKLQTIFDTLPPMRYQETMSAIRTWVQQSETKLSIPQLSVTDY
                     EIMEQRLGELQALQSSLQEQQSGLYLYSTTVKEMSKKAPSEISRKYQSEFEIEGRWK
                     KLSSQLVEHCQKLEEQMNKLRKIQNHIQTLKKWMAEVDVFLKEEWPALGDSEILKQL
                     KQCRLLVSDIQTIQPSLNSVNEGGQKIKNAEPEFASRLETELKELNTOWDHMCCQVY
                     ARKEALKGGLEKTVSLQKDLSEMHEWMTQAEEYLERDFEYKTPDELQKAFEEMKRAK
                     EEAQQKEAKVKLLTESVNSVIAQAPPVAQEALKLKELTLLTTNYQWICTRLNGKCKTLE
                     EVWACWHELLSVYLEKANKWLNEVEFKLKTTENIPGGAEEISEVLDSLENLMRHSEDNP
                     NQIRILAQTILTDGGVMDELINEELETFNSRWRELHEEAVRQKLLEQSIQSAQETENS
                     LHLIQESLTFIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLEEMKKHNQGKEA
                     AQRVLSQIDVAQKLQDVSMKFRLFQKPANFEQRLQESKMLLDEVKMHLPALETKSVE
                     QEVVQSQLNHCVNLYKSLSEVKSEVEMVIKTGRQIVQKKQTENPKELDERVTALKLHY
                     NELGAKVTERKQQLEKCLKLSRKMRKEMNVJTEWLAATDMELTKRSAVEGMPSNLDSE
                     EWLNLLSYQKHMETEDQNVDHITKWIIQADTLLDESEKKKPQQKEDVLKRLKAELND
                     IRPKVDSTRDQAANLMANHGDRCRKLVEPQISELNHRFAAISHRIKTGKASIPLKELE
                     QFNSDIQKLLEPLEAEIQQGVNLKEEDFNKDMNEDNEGTVKELLQRGDNLCQRITDER
```

```
Db    8956  ARAATTGAACCTGCCACTCCGCTGACTGGCGAGAGAAAAATAGATGAGACCCTTGAAAGACT  9015

Qy     301  CCAGGAACTTCAAGAGGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT   360
            ||||||||||||||||||||| ||||||||       |||||||||||||||||||||
Db    9016  CCAGGAACTTCAAGAGGGCCACGGACGGATGAGCTGCGCCAAGCTGAGGTGAT         9075

Qy     361  CAAGGGATCCTCTGCAGCCGCCGTGGGCGATCTCCTCATTGACTTCTCCAAGATCACCTCGA  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9076  CAAGGGATCCTCGGCAGCCGCCGTGGGCGATCTCCTTGACTTCTCCAAGATCACCTCGA   9135

Qy     421  GAAAGTCAAGGCACTTCGAGGAGAAATTGCCGCCTCTGAAAGAGAGAAGTGAGCCACGTCAA  480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9136  GAAAGTCAAGGCACTTCGAGGAGGAGAAATTGCGCCTCTGAAAGAGAGAACGTGAGCCACGTCAA  9195

Qy     481  TGACTTGCTCGCGAGCTTACCACTTTGGGCATTCAGCTCTCACGTATAACCTCAGCAC   540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9196  TGACTTGCTCGCGAGCTTACCACTTTGGGCATTCAGCTCTTCACCGTATAACCTCAGCAC   9255

Qy     541  TCTGAAGACCTGAACACCAGATGGAAGCTTCTGCAGTGGCCGCGTCGAGGACCGAGTCAG   600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9256  TCTGAAGACCTGAACACCAGATGGAAGCTTCTGCAGTGGCCGCGTCGAGGACCGAGTCAG   9315

Qy     601  GCAGCTGCATGAGAGCCCACAGGGACTTTGGTCCAGCACTTCTTCTTCCACGTC       660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9316  GCAGCTGCATGAGAGCCCACAGGGACTTTGGTCCAGCACTTTCTTCTTCCACGTC       9375

Qy     661  TGTCCAGGGTCCCTGGGAGGAGCAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9376  TGTCCAGGGTCCCTGGGAGGAGCAGCCATCTCGCCAAACAAAGTGCCCCTACTATATCAACCA  9435

Qy     721  CGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACACAGAGCTCACCAGTCTTTAGC   780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9436  CGAGACTCAAACAACTTGCTGGGACCATCCCAAATGACACAGAGCTCACCAGTCTTTAGC   9495

Qy     781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9496  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA   9555

Qy     841  GAAGGCCCTTGCTTGGATCTCTTGAGCCGTGTCAGCTGCATGTGATGCCCTTGGACCAGCA  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9556  GAAGGCCCTTGCTTGGATCTCTTGAGCCGTGTCAGCTGCATGTGATGCCCTTGGACCAGCA  9615

Qy     901  CAACCTCAAGCAAAATGACCAGCCCATGAACATCCTGCAGATTATTAATTGTTGACCAC    960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9616  CAACCTCAAGCAAAATGACCAGCCCATGAACATCCTGCAGATTATTAATTGTTGACCAC    9675

Qy     961  TATTTATGACCGCCTGGAGCAAGAGCACAACAACAATTGGTCAACGTCCCCTCTCGCGTGGA  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9676  TATTTATGACCGCCTGGAGCAAGAGCACAACAACAATTGGTCAACGTCCCCTCTCGCGTGGA  9735

Qy    1021  TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACAACAGGACAGGAGGATCCTGT  1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9736  TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACAACAGGACAGGAGGATCCTGT  9795

Qy    1081  CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGACACATTGGAAGACAAAGTACAC  1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9796  CCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGACACATTGGAAGACAAGTACAC   9855

Qy    1141  ATACCTTTTCAAGCAAGTGCAAGTTCTATCCAAATTCCAAGACAGATTTTGTGACCCAGGCCCT  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9856  ATACCTTTTCAAGCAAGTGCAAGTTCTATCCAAATTCCAAGACAGTTTTGTGACCCAGGCCCT  9915

Qy    1201  CCTTCTGCATGATTCTATCCAAATTCCAAGACACAGTTGCATCCTTGGGGG         1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9916  CCTTCTGCATGATTCTATCCAAATTCCAAGACACAGTTGGCATCCTTGGGGG        9975

Qy    1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG                      1301
            |||||||||||||||||||||||||||||||||||||||
Db    9976  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG                      10016
```

RESULT 8
HSDMDR

```
Qy   541  TCTGGAAGACCTGAACACCAGACCTGGCCTGCAGGTGGCCTGCAGGACCGAGTCAG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9146  TCTGGAAGACCTGAACACCAGACCTGGCCTGCAGGTGGCCTGCAGGACCGAGTCAG  9205

Qy   601  GCAGCTGCATGAAGCCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9206  GCAGCTGCATGAAGCCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGTC  9265

Qy   661  TGTCCAGGTGCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9266  TGTCCAGGTGCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA  9325

Qy   721  CGAGACTCAAACAACTGCTGGGACCATCCCAAAATGACAGAGCTTCACCAGTCTTTAGC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9326  CGAGACTCAAACAACTGCTGGGACCATCCCAAAATGACAGAGCTTCACCAGTCTTTAGC  9385

Qy   781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCTGCATGTGATGCCGAGACTGCA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9386  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCTGCATGTGATGCCGAGACTGCA  9445

Qy   841  GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGACCAGCA   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9446  GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGACCAGCA  9505

Qy   901  CAACCTCAAGCAAAATGACCAGCCCATGAGATCCTGCAGATTATTAATTGTTTGACCAC   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9506  CAACCTCAAGCAAAATGACCAGCCCATGAGATCCTGCAGATTATTAATTGTTTGACCAC  9565

Qy   961  TATTTATGACCCGCCCTGAGCAAGAGCAACAACAATTTGGTCAACGTCCCTCGCGTGGA  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9566  TATTTATGACCCGCCCTGAGCAAGAGCAACAACAATTTGGTCAACGTCCCTCGCGTGGA  9625

Qy  1021  TATGTGTCTGAACTGGCTGCTGCTGAATGTTTATGATACGGACGAACAGCAGGAGATCCGTGT  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9626  TATGTGTCTGAACTGGCTGCTGCTGAATGTTTATGATACGGACGAACAGCAGGAGATCCGTGT  9685

Qy  1081  CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAAGCACATTTGGAAGACAAGTACAG  1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9686  CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAAGCACATTTGGAAGACAAGTACAG  9745

Qy  1141  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCCGCAGGCTGGGCCT  1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9746  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAGCCGCAGGCTGGGCCT  9805

Qy  1201  CCTTCTGCATGATTCATCCAAATTCAAGACAGTTGGGTGGTGCATCCTTTGGGGG  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9806  CCTTCTGCATGATTCATCCAAATTCAAGACAGTTGGGTGGTGCATCCTTTGGGGG  9865

Qy  1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301
          |||||||||||||||||||||||||||||||||||||||||
Db  9866  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  9906
```

Protein sequence:

```
KSEEIKIKQLLQTKHNALKDLRSQRRKKALEISHQWYQYKRQADDLLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWRE
IESKRAQFRRLNFAQIHTVREETMMVMTEDMPLEISVVPSTYLTEITHVSQALLEVEQ
LLNAPDLCAKDFEDLFKQBESLKNIKDSLQQSSGRIDIIHSKKTAALQSATPVERVKL
QEALSQLDFQWEKVNKMYKDRQGRFDRSVEKWRRFHYDIKIFNQWLTEAEQFLRKTQI
PENWEHAKYKWYLKELQDGIGQRQTVVRTLNATGEEIIQQSSKTDASILQEKLGSINL
RWQEVCKKLSDRKKRLBEQXNILSEFQRDLNEFVLWLEEADNIASIPLEPGKEQQLKE
KLEQVKLLVEPLRQGIIKQLNETGGPVLVSAPISPEQDKLENLKQWIKVS
RALPEKQGBEIBAQIKDLQGLEKKLEDLEEQLNHLLLWLSPIRNQLBIYNQPNQEGPFD
VKETBIAVQAKQPDVEBIILSKGQHLYKEKPATQVKKRLEDLSSEWKAVNRLLQELRA
KQPDLAPGLJTIGASPTQVTLVTQPVVTKETAISKLEMPSSSLMLEVPALADFNRAWT
ELTDWLSLLDQVIKSQRVMVGDLBDINEMIIKQKATMQDLEQRRPQLEELITAAQNLK
NKTSNQEARTIITDRIERIQNQWDEVQEHLQNRRQQLNEMLKDSTQWLEAKEBABQVL
GQARAKLESWKEGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSAD
DTRKVHMITENINASWRSIHKRVSEREALEETHRLLQQFPLDLEKFLAWLTEAETTA
NVLQDATRKERLLEDSKGVKELMKQWQDLQGEIEAHTDVYHNLDENSQKILRSLEGSD
DAVLLQRRLDNMNFKKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELLVWLQLKDDEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKTKEPVIMSTLETVRIFLITBQPLEGLEKLYQ
EPRELPEERAQNVTRLLRKQAEEVNTEWEKLNHLHSADWQRKIDETLERLQELQEATD
ELDIKLRQAEVIKGSWQFVGDLLIDSLQDHLEXVKALRGEIAPLKENVSHVNDLARQL
TTLGIQLSFPNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVQGP
WERAISPNKVPYYINHETQTCWDHPKMTELYQSLADINNVRFSAYRTAMKLRRLQKA
LCDLLSLSAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHNNLVNVPLCVD
MCLNWLLNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRL
GLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFPANNKPEIBAALFDWMRLEBPQSM
VWLPVLHRVAAAETAKHQAKCNICKECPIIGFYRSLKHPNYDICQSCFPSGRVAKGH
KMHYPMVEYCTPTISGEDVRDFAKVLKNKFRTKRYFAKPRMGYLPVQTVLEGDNMET
PVTLINFWPVDSAPASSPQLSHDDTHSRIEHYASRLAEMENSNGSYLNDSISPNESID
DEHLLIQHYCQSLNQDSPLSQPRSPAQILSLESEERGELERILADLEBEENRLQAEY
DRLKQOHEHKGLSPLPSPPEMMPTSPQSPRDAELIAEAKLLRQHGKGLEEARMQILEDH
NKQLESQLHRIRQLLEQPQAEAKVNGTTVSSPSTSLQRSDSSOPMLLRVVGSQTSDSM
GEBDLLSPPQDTSTGLEEVMEQLNNSFPSSRGRNTPGKPMREDTM"
```

```
ORIGIN
Query Match        100.0%;   Score 1301;   DB 9;   Length 12446;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1301;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1  CGACTTTCCAGCAGTTCAAGAACGACGAGATGTACATAGGCCCTTCAAGAGGGAATTGAA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8606  CGACTTTCCAGCAGTTCAAGAACGACGAGATGTACATAGGCCCTTCAAGAGGGAATTGAA  8665

Qy    61  AACTAAAGAACCTGTAATCATGAGACTACTCTTGAGACTGTACGAACATATTTCTGACAGAGCA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8666  AACTAAAGAACCTGTAATCATGAGACTACTCTTGAGACTGTACGAACATATTTCTGACAGAGCA  8725

Qy   121  GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCTGCCTCCTCCGAGGAGAG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8726  GCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCTGCCTCCTCCGAGGAGAG  8785

Qy   181  AGCCCAGAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8786  AGCCCAGAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGA  8845

Qy   241  AAAATTGAACCTGCACTCCGCTGACTGGCACAGCAGAGAAAAATGAGACCCTTGAAAGACT   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8846  AAAATTGAACCTGCACTCCGCTGACTGGCACAGCAGAGAAAAATAGATGAGACCCTTGAAAGACT  8905

Qy   301  CCAGGAACTTCAAGAGGCCACGGCCCGTGGCCCGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGCTGAT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8906  CCAGGAACTTCAAGAGGCCACGGCCCACGGTGGCCCGATGAGCTGACCTCAAGCTGCGCCAAGCTGAGCTGAT  8965

Qy   361  CAAGGGGATCCTGCAGCCCGTGGGCGAATCTCTCATTGACTCTCCAAGATCACCTCGA   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8966  CAAGGGGATCCTGCAGCCCGTGGGCGAATCTCCTCATTGACTCTCCAAGATCACCTCGA  9025

Qy   421  GAAAGTCAAGGACTTCAAGAGGACACTTCGAGGAGAAATTGCGCCTCGAAAGAACGTGAGCCACGTCAA   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9026  GAAAGTCAAGGACTTCAAGAGGACACTTCGAGGAGAAATTGCGCCTCGAAAGAACGTGAGCCACGTCAA  9085

Qy   481  TGACCTTGCTCGCCAGCTTACCACTTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9086  TGACCTTGCTCGCCAGCTTACCACTTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC  9145
```

ORIGIN

```
       /mol_type="unassigned DNA"
       /db_xref="taxon:9606"
       /note="EMBL/GenBank Accession No. M18533"

Query Match      100.0%;    Score 1301;    DB 6;    Length 13957;
Best Local Similarity  100.0%;    Pred. No.0;
Matches 1301;    Conservative  0;    Mismatches  0;    Indels  0;    Gaps  0;
```

```
Qy      1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAA    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8716 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAA  8775

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGACTCTTGACAGAGCA   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8776 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGACTCTTGACAGAGCA  8835

Qy    121 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGCGCCAGAGAGCTGCCTCCTGAGGAGAG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8836 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGCGCCAGAGAGCTGCCTCCTGAGGAGAG  8895

Qy    181 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCACGAAAGCAGCTGCCTCACGAAAGACT   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8896 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCACGAAAGCAGCTGCCTCACGAAAGACT  8955

Qy    241 AAAATTGAACCTGCACTCCCGCTGACTCGGCGCAGAGAAAATAGATGAGACCCTTGAAAGACT   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8956 AAAATTGAACCTGCACTCCCGCTGACTCGGCGCAGAGAAAATAGATGAGACCCTTGAAAGACT  9015

Qy    301 CCAGGAACTTCAAGAGGCCACGCGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9016 CCAGGAACTTCAAGAGGCCACGCGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT  9075

Qy    361 CAAGGATCCTGGCCAGCTACCACTTTGGGCATTCAGTCTCCATTGGACTCTCCAAGAATCACCTCGA   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9076 CAAGGATCCTGGCCAGCTACCACTTTGGGCATTCAGTCTCCATTGGACTCTCCAAGAATCACCTCGA  9135

Qy    421 GAAAGTCAAGGCACTTCGAGAGAATTGGCCTCTGAAAAGAACGTGAGCCACGTCAA   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9136 GAAAGTCAAGGCACTTCGAGAGAATTGGCCTCTGAAAAGAACGTGAGCCACGTCAA  9195

Qy    481 TGACCTGCTCGCCAGCTACCACTTTGGGCATTCAGTCTCCAAGAATCACCTCGA   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9196 TGACCTGCTCGCCAGCTACCACTTTGGGCATTCAGTCTCCAAGAATCACCTCGA  9255

Qy    541 TCTGGAAGACTGAACACCAGACCAGATGGAAGCTTCTGCAGGTGCCGTGAGGACCGAGTCAG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9256 TCTGGAAGACTGAACACCAGACCAGATGGAAGCTTCTGCAGGTGCCGTGAGGACCGAGTCAG  9315

Qy    601 GCAGCTGCATGAAGCCCACAGGGACTTGGTCCAGCACTTTCTTCCACGTC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9316 GCAGCTGCATGAAGCCCACAGGGACTTGGTCCAGCACTTTCTTCCACGTC  9375

Qy    661 TGTCCAGGGTCCTGGGAGAGCCATCTGCCCAAACAAAGTGCCACTACTATATCAACCA   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9376 TGTCCAGGGTCCTGGGAGAGCCATCTGCCCAAACAAAGTGCCACTACTATATCAACCA  9435

Qy    721 CGAGACTCAAACAACTGCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9436 CGAGACTCAAACAACTGCTGGACCATCCAAAATGACAGAGCTCTACCAGTCTTTAGC  9495

Qy    781 TGACCTGAATAATGTCAGATTCTCAGATTCTCAGCTTTATAGGACTGCCATGAAAACTCCGAGAACTGCA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9496 TGACCTGAATAATGTCAGATTCTCAGATTCTCAGCTTTATAGGACTGCCATGAAAACTCCGAGAACTGCA  9555

Qy    841 GAAGGCAATGTCTTGAGCTTGGTTGATGCTTCAGCTGCGCTGCGCTGAGCGTCATGATGCTTGACCAGCA   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9556 GAAGGCAATGTCTTGAGCTTGGTTGATGCTTCAGCTGCGCTGCGCTGAGCGTCATGATGCTTGACCAGCA  9615

Qy    901 CAACCTCAAGCAAAAATGAACCTCGCACTCCGCTGACTATCCTGCAGATATCCTCGCAGATTATTAATTGTTTGACCAC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9616 CAACCTCAAGCAAAAATGAACCTCGCACTCCGCTGACTATCCTGCAGATATCCTCGCAGATTATTAATTGTTTGACCAC  9675
```

```
Qy    961 TATTTATGACCGCCTGGAGCAAGAGCACAACACAATTTGGTCAACGTCCCTCTGCGTGGA   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9676 TATTTATGACCGCCTGGAGCAAGAGCACAACACAATTTGGTCAACGTCCCTCTGCGTGGA  9735

Qy   1021 TATGTGTCTGAACTGGCTGCTGATGTTTATGATACGGGACGAACAGGGGGATCCGTGT   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9736 TATGTGTCTGAACTGGCTGCTGATGTTTATGATACGGGACGAACAGGGGGATCCGTGT  9795

Qy   1081 CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9796 CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG  9855

Qy   1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGGCAGGCTGGGCCT   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9856 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGGGCAGGCTGGGCCT  9915

Qy   1201 CCTTCTGCGATGATTCATCCAAATTCCAAGACAGTGGGTGAGTTGCATCCTTGGGGG   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9916 CCTTCTGCGATGATTCATCCAAATTCCAAGACAGTGGGTGAGTTGCATCCTTTGGGGG  9975

Qy   1261 CAGTAACATTGAGCCAAGTGTCCCGAGCTGCTTCCAATTTG   1301
          |||||||||||||||||||||||||||||||||||||||||
Db   9976 CAGTAACATTGAGCCAAGTGTCCCGAGCTGCTTCCAATTTG  10016
```

```
RESULT 10
AX538581
LOCUS       AX538581         13957 bp    DNA    linear    PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO0229056.
ACCESSION   AX538581
VERSION     AX538581.1  GI:25271086
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE     Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 1 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match      100.0%;    Score 1301;    DB 6;    Length 13957;
Best Local Similarity  100.0%;    Pred. No. 0;
Matches 1301;    Conservative  0;    Mismatches  0;    Indels  0;    Gaps  0;
```

```
Qy      1 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAA    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8716 CGACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGGCCTTCAAGAGGGAATTGAA  8775

Qy     61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGACTCTTGACAGAGCA   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8776 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGACTCTTGACAGAGCA  8835

Qy    121 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGCGCCAGAGAGCTGCCTCCTGAGGAGAG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8836 GCCTTTGGAAGGACTAGAGAAACTTCTACCAGGAGCGCCAGAGAGCTGCCTCCTGAGGAGAG  8895

Qy    181 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCACGAAAGCAGCTGCCTCACGAAAGACT   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8896 AGCCCAGAATGTCACTCGGCTTCACTCGGCTTCACGAAAGCAGCTGCCTCACGAAAGACT  8955

Qy    241 AAAATTGAACCTGCACTCCGCTGACTCGCCCGCTGGACATGAGCTGCGCGCAGAGAAAATAGATGAGACCCTTGAAAGACT   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8956 AAAATTGAACCTGCACTCCGCTGACTCGCCCGCTGGACATGAGCTGCGCGCAGAGAAAATAGATGAGACCCTTGAAAGACT  9015

Qy    301 CCAGGAACTTCAAGAGGCCACGCGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9016 CCAGGAACTTCAAGAGGCCACGCGATGAGCTGGACCTCAAGCTGCGCCAAGCTGAGGTGAT  9075
```

```
Db  9016 CCAGGAACTTCAAGAGGCCACGGATGAGCTCGGACCTCCTCATTGACTCTGCGCCAAGCTGAGGTGAT 9075
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361 CAGGGGATCCTGCAGCCACGGATCTCCAGCCACGGGCCGCATCCTCCTCCAAGATCACCTCGA 420
Db  9076 CAGGGGATCCTGCAGCCACGGATCTCCAGCCACGGGCCGGATCTCCTCCAAGATCACCTCGA 9135

Qy   421 GAAAGTCAAGGCACTTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 480
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  9136 GAAAGTCAAGGCACTTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAA 9195

Qy   481 TGACCTTGCTGCCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTTCAGCAC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9196 TGACCTTGCTGCCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTTCAGCAC 9255

Qy   541 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGGAGGACCGAGTCAG 600
         |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  9256 TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGGCCCGTCGGAGGACCGAGTCAG 9315

Qy   601 GCAGCTGATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGACACTTCTTTCCACGTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9316 GCAGCTGATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAGACACTTCTTTCCACGTC 9375

Qy   661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9376 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 9435

Qy   721 CGAGACTCAAACAACAACTTGCTGGGACTCTGGATCTCTTGCTTGAGCCTGTCAGTCTTTAGC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9436 CGAGACTCAAACAACAACTTGCTGGGACTCTGGATCTCTTGCTTGAGCCTGTCAGTCTTTAGC 9495

Qy   781 TGACCTGAATAATGTCAGATTCTCAGCTTAGGACTCCATGAAACTCCGAAGACTGCA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9496 TGACCTGAATAATGTCAGATTCTCAGCTTAGGACTCCATGAAACTCCGAAGACTGCA 9555

Qy   841 GAAGGCCCTTGCTTGGATCTCTTGAGCCTGTGCATGGCGTGCATGGACCAGCA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9556 GAAGGCCCTTGCTTGGATCTCTTGAGCCTGTGCATGGCGTGCATGGACCAGCA 9615

Qy   901 CAACCTCAAGCAAAATGACCAGCCATGGATATCCTGCAGATATCCTGGTTGTTTGACCAC 960
         |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  9616 CAACCTCAAGCAAAATGACCAGCCATGGATATCCTGCAGATTAATTGTTTGACCAC 9675

Qy   961 TATTTATGACCGGCCTGGAGCAAGAGACAACAATTGGTCAACGTGTCCCTCTCGCTGGA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9676 TATTTATGACCGGCCTGGAGCAAGAGACAACAATTGGTCAACGTGTCCCTCTCGCTGGA 9735

Qy  1021 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACAGAAGGGATCCGTGT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  9736 TATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACAGAAGGAGGATCCGTGT 9795

Qy  1081 CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACACAGTTGGAAGCACAAGTACAG 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9796 CCTGTCTTTTAAAACTGGCATCATTCCCTGTGTAAAGCACACAGTTGGAAGCACAAGTACAG 9855

Qy  1141 ATACCTTTCAAGCAAGTGGCAAGTTGTGACCGCGCAGGCTGGGCCT 1200
         |||||||||||||||||||||||||||||||||||| ||||||||
Db  9856 ATACCTTTCAAGCAAGTGGCAAGTTGTGACCGCGCGGCTGGGCCT 9915

Qy  1201 CCTTCTCGATGATTCTATCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9916 CCTTCTCGATGATTCTATCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGG 9975

Qy  1261 CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG 1301
         ||||||||||||||||||||||||||||||||||||||||
Db  9976 CAGTAAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTG 10016
```

RESULT 11

HUMDYS

```
LOCUS       HUMDYS     13957 bp    mRNA    linear   PRI 25-MAY-2000
DEFINITION  Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION   M18533 M17154 M18026 M20250
VERSION     M18533.1  GI:181856
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1699)
  AUTHORS   Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and
            Kunkel,L.M.
  TITLE     Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
            preliminary genomic organization of the DMD gene in normal and
            affected individuals
  JOURNAL   Cell 50 (3), 509-517 (1987)
  MEDLINE   87273512
   PUBMED   3607877
REFERENCE   2  (bases 1678 to 3830)
  AUTHORS   Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
  TITLE     Conservation of the Duchenne muscular dystrophy gene in mice and
            humans
  JOURNAL   Science 238 (4825), 347-350 (1987)
  MEDLINE   88018015
   PUBMED   3659917
REFERENCE   3  (bases 1 to 13957)
  AUTHORS   Koenig,M., Monaco,A.P. and Kunkel,L.M.
  TITLE     The complete sequence of dystrophin predicts a rod-shaped
            cytoskeletal protein
  JOURNAL   Cell 53 (2), 219-226 (1988)
  MEDLINE   88194521
   PUBMED   3282674
COMMENT     On May 25, 2000 this sequence version replaced gi:340693.
            Draft entry and computer-readable sequence kindly provided by
            M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
            determined by the size of the deleted DNA segment. Deletions found
            in different patients were from positions 302-2200, 473-1168,
            1691-1810, and 1169-3011.
FEATURES             Location/Qualifiers
     source          1..13957
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xp21.3-p21.1"
                     /tissue_type="muscle"
                     /dev_stage="fetus"
                     /note="G00-119-850"
     gene            1..13957
                     /gene="DMD"
     CDS             209..11266
                     /gene="DMD"
                     /codon_start=1
                     /product="dystrophin"
                     /protein_id="AAA53189.1"
                     /db_xref="GI:181857"
                     /translation="MLWWEEVEDCYEREDVQKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVHALNNVNKALRVLQNNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIIHHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDV
                     DTTYPDKKSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEEHFQLNHQMHYSQ
                     QITVSLAQGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFFSQHLRAPEDKSFGSSL
                     MESEVNLDRYQTALEEVLSWLLSAEDTLQAQGEISNDVEVVKDQFHTHEGYMMDLTAH
                     QGRVGNILQLGSKLIGTGKLSEDEETEVQGMNLLNSRWECLRVASMEKQSNLHRVLM
                     DLQNQKLKELNDWLTKTEERTRKMEEEPLGPDLEDLKRQVQQHKVLQEDLEQEQVRVN
                     SLTHMVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWVLLQDILLKWQRLTE
                     EQCLFSAWLSEKEDAVNKIHTTGFKDQNEMLSSLQKLAVLKADLEKKKQSMGKLYSLK
                     QDLLSTLKNKSVTQKTEAWLDNFARCWDNLVQKLEKSTAQISCAVTTTQPSLTQTTVM
                     ETVTVTTREQILVKHAQELPPPPQKRQITVDSEIRKLRDVDITELHSWITRSEA
                     VLQSPFAIPRKEGNFSDLKEVNAIERKAEKFPRKLQDASRSAQALVEQMVNEGVNA
                     DSIKQASEQLNSRWIEFCQLLSERLNWLEYQNNIIAFYNQLQQLEQMTTTAENWLKIQ
                     PTPSEPTAIKSQLKICKDKVNRLSGLQPQIERLKIQSIALKEKGQGPMFLDADFVAF
                     TNHFKVFSVQVARKELQTIFDTLPPMYQETMSAIRTWVQQSETKLSIQQSVTDY
                     EIMEQRLGELQALQSSLEQQGSLYYLSTTVKEMSKKAPSEISRKYQSEFEEIEGRWK
                     KLSSQLVEKCGKLEEQMNKLRKIQNHIQTLKKWMAEVDVFLKEEWPALGDSELKKQL
                     KQCRLLVSDIQTIQPSLNSVNEGGQKIKNEAEPFASRLETELKELNTQWDHMCQQVY"
```

```
Qy   361  CAAGGGATCCTGCCAGCCGTGGCCATCTCCTGGCCATCTCTCTCCTCCAAGATCACCTCGA  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9076  CAAGGGATCCTGCCAGCCGTGGCCGGATCTCCTGGCCATCTCTCTCCTCCAGATCACCTCGA  9135

Qy   421  GAAAGTCAAGGCACTTGAGAGAAATGCGGCCTCTGAAAGAGAAGTGAGCCACGTCAA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9136  GAAAGTCAAGGCACTTGAGGAGAAATGCGGCCTCTGAAAGAGAAGTGAGCCACGTCAA  9195

Qy   481  TGACCTTGCTGCCAGTTACCACTTTGGGATTCAGCTCTCCACCGTATAACCTCAGCAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9196  TGACCTTGCTGCCAGCTTACCACTTTGGGCATTCAGCTCTCCACCGTATAACCTCAGCAC  9255

Qy   541  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGAGGACCGAGTCAG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9256  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTCGAGGACCGAGTCAG  9315

Qy   601  GCAGCTGCATGAGAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9316  GCAGCTGCATGAGAGCCCACAGGGACTTTGGTCCAGCATCTCAGCACTTCTTTCCACGTC  9375

Qy   661  TGTCCAGGGTCCCTGGGAGAGCCATCTCGGCCAAACAAAGTGCCCTACTATATCAACCA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9376  TGTCCAGGGTCCCTGGGAGAGCCATCTCGGCCAAACAAAGTGCCCTACTATATCAACCA  9435

Qy   721  CGAGACTCAAACAACTGCTGGGACCATCCCCAAAATGACAGAGCTCTCCACCGTCTTTAGC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9436  CGAGACTCAAACAACTGCTGGGACCATCCCCAAAATGACAGAGCTCTCCACCGTCTTTAGC  9495

Qy   781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9496  TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA  9555

Qy   841  GAAGGCCCTTTGCTTGGATCCTTGAGCCTGTCAGCTGCATGGATGCCTTGGACCAGCA  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9556  GAAGGCCCTTTGCTTGGATCCTTGAGCCTGTCAGCTGCATGGATGCCTTGGACCAGCA  9615

Qy   901  CAACCTCAAGCAAAATGACCAGCCCATGGACATCCTGCAGATTATTAATTGTTGACCAC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9616  CAACCTCAAGCAAAATGACCAGCCCATGGACATCCTGCAGATTATTAATTGTTGACCAC  9675

Qy   961  TATTTATGACCGCCTGGACAAGAGCAAGAGCAACAATTGGTCAACGTCCCTCTGCGTGGA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9676  TATTTATGACCGCCTGGACAAGAGCAAGAGCAACAATTGGTCAACGTCCCTCTGCGTGGA  9735

Qy  1021  TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGACGAACAGGAGGATCCGTGT  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9736  TATGTGTCTGAACTGGCTGCTGAATGTTATGATACGGACGAACAGGAGGATCCGTGT  9795

Qy  1081  CCTGTCTTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9796  CCTGTCTTTTAAAACTGGCATCATTCCCGTGTAAAGCACATTTGGAAGACAAGTACAG  9855

Qy  1141  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCAGGCTGGGCCT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9856  ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCAGGCTGGGCCT  9915

Qy  1201  CCTTCTGCATGATTCTATCCAAATTCCAAGAGAGCTGCCTCCTGGAGGAGAG  1260
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  9916  CCTTCTGCATGATTCTATCCAAATTCCAAGAGAGCTGCCTCCTGGAGGAGAG  9975

Qy  1261  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  1301
          |||||||||||||||||||||||||||||||||||||||||
Db  9976  CAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG  10016
```

RESULT 12
AR220819
```
LOCUS       AR220819            13977 bp    DNA       linear    PAT 26-SEP-2002
DEFINITION  Sequence 60 from patent US 6426186.
ACCESSION   AR220819
VERSION     AR220819.1  GI:23327696
KEYWORDS    .
```

ORIGIN

```
Query Match   100.0%;  Score 1301;  DB 9;  Length 13957;
Best Local Similarity 100.0%;
Matches 1301;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1  CGACTTCCAGCAGTTCAAGAACAGAACGATGTACATAGGGCCTTCAAGACGGGAATTGAA  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8716  CGACTTCCAGCAGTTCAAGAACAGAACGATGTACATAGGGCCTTCAAGACGGGAATTGAA  8775

Qy   61  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACACAGACA  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8776  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACACAGACA  8835

Qy  121  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAAAGCAGGCTGAGGAGGTCAATACTGAGTTGGGA  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8836  GCCTTGGAAGGACTAGAGAAACTCTACCAGGAAAGCAGGCTGAGGAGGTCAATACTGAGTTGGGA  8895

Qy  181  AGCCCAGAGTCACTCGGCTTCACGTCGCTGAGAAAGCAGGAGGAGGTCAATACTGAGTGGGA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8896  AGCCCAGAGTCACTCGGCTTCACGTCGCTGAGAAAGCAGGAGGAGGTCAATACTGAGTGGGA  8955

Qy  241  AAAATTGAACCTGACTCGGATGAGCGCGCCAAGTGCGCGCCAAGCTGAAAGACT  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8956  AAAATTGAACCTGACTCGGATGAGCGCGCCAAGTGCGCGCCAAGCTGAAAGACT  9015

Qy  301  CCAGGAACTTCAAGAGGCCACGGAGCCGCGGTGGACTCCAAGCTGCGCCAAGCTGCGCCAAGCTGGA  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9016  CCAGGAACTTCAAGAGGCCACGGAGCCGCGGTGGACTCCAAGCTGCGCCAAGCTGCGCCAAGCTGGA  9075
```

Protein sequence:

ARKEALKGGLEKTVSLQKDLSEMHEWMTQAEEEYLERDFPEYKTPDELQKAVEEMKRAK
EEAQQKEAKVKLLTESVNSVIAQAPPVAQEALKKELETLLTTNVQWLCTRLNGKCKTLE
EVWACWHELLSYLEKANKWLNEVEFFKLKTTENIPGGAEEISEVULDSLENLMRHSEDNP
NQIRILAQTLTDGGVMDELINEELETPNSRWRELHEEAVRRQKLLEQSLQSAQETEKS
LHLIQESLTFIDKQLAAYIADKVDAAQMPQEAQKIQSDLTSHEISLEEMKKHNQGKEA
AQRVLSQIDVAQKLQDVSMKFRLFQKPANFELRIQESKMILDEVKMHLPALETKSVE
QEVVQSLNHCVNLYKSLSEVYKSEVEMVIKTGRQIVQKKQTENPKELDERVTAIKLHY
NELGAKVTERKQQLEKCLKLSRKMRKEMNVLTEWLAATDMELTKRSAVEGMPSNLDSE
VAWGKATQKEIEKQKVHLKSITEVGBALKVLGKKETLVEDKLSLLNSNWIAVTSRAE
EWLNLILEYQKMETFPDQNVDHITKNIIQADTLLDDSEKKKPQQYEDVLKRLKAELND
IRPKVDSTRDQAANLMANRGDHCRKLVEPQISELNHRFAAISHRIKTGKASIPLKELE
QFNSDIQKLLEPLEEAEIQQGVNLKEEDFNKDMNEDNEGTVKELLQRGDNLQQRITDER
KREEIKIKQQLLQTKHNALKDLRSQRRKKALEISHQWYQKRQADDLLKCLDDIEKKL
ASLPEPRDERKIKEIDRELQKKEELNAVRRQAEGLSEDGAAMAVEPTQIQLSKRWRE
IESKFAQFRRLNFAQIHTVREETMMVMTBDMPLEISYVPSTYLTEITHVSQALLEVEQ
LLNAPDLCAKDFEDLFKQEEBSLKNIKDSLLQQSSGRIDIIHSKKTPAALQSATPVERVKL
QEALSQLDFQWEKVNKMYKDRQGRFDRSVEKWRFPHYDIKIFNQWLTEAEQFLRKTQI
PENWEHAKYKNYLKELQDGIQGRQTVVRTLNATGBEIIQQSSKTDASIIQEKLGSLNL
RWQEVCKQLSDRKKRALEEQKNILSEFQRDLNBFVLWLEBADNIASIPLEPGKEQQLKE
KLEQVKLLVEELPLRQGLKQLNETGGPVLVSAPISPEEQDKLENKLKQTNQGMIKVS
RALPEKQGEIEAQIKDLQGLEKKLEDLEBQLNHLLMLSPIRNQLEIYNQPQEGPFD
VQETEIAVQAKQPDVBEILSKGQHLYKEKPATQVPVVTKTAISKLEMPSSLMLEVPALADFNRAWT
KQPDLAPGLITTIGASPTQTVLIVTQPVVTKTAISKLEMPSSLMLEVPALADFNRAWT
ELTDWLSLLDQVIKSQRVMVGDLEDLNEMIIKQKATMQDLEQRRPQLEBLIITAAQNLK
NKTSNQEARTIITDRIERIQNQWDEVQEHLQNRRQQLNEMLKDSTQWLEAKEEABQVL
GQARAKLESWKEGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLAIKLLRDYSAD
DTRKVHMITENINASWRSIHKCRVSEREAALEETHRLLQQFPLDLDEKFLAWLTEAETTA
NVLQDATRKERLLEDSKGVKELMKGWQDLQGEIRAHTDVYHNLDENSQKILRSLEGSD
DAVLLQRRLDNMPFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELIVWLQLKDDEL
SRQAPIGGDFPAVQKQNDVHRAFKRELKTEKSPIVIMSTLBTVRIFITEQPLEGLEKLYQ
EPRELPBERAQNVTRLLRKQAEBVNTEWEKLNLHSADWQRKIDETLERLQELQBATD
ELDKLRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEIAPLKENVSHVNDLARQL
TTLIGIQSPYNLSTLEDINTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVQGP
WERAISPNKVPYYINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKA
LCLDLSLSAACDALDQHNLKQNDQPMDIIQIINCLTTIYDRLEQBHNNLVNPLCVD
MCLNWLLNVDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRL
GLLLHIDSIQIPRQLGBVASFGGSNIEPSVRSCFQFANNKPEEAALFLDWMRLEPQSM
VWLPLHRVAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDICQSCFFSGRVAKGH
KMHYPMVEYCTPTTSGEDVRDFAKVLKNKFRTKKYFAKHPRMGVLPVQTVLEGDNMET
PVTLINFWPVDSAPASSPQLSHDDTTHSRIEHYASRLAEMENSNGSYLNDSISPNESID
DEHLLIQHYCQSLNQDSPLSQPRSPAQILISLESEERGELERIIADLEEENRNLQABY
DRLKQOHEHKGLSPLPSPPEMMPTSPQSPRDAELIAEAKLLRQHKGRLEARMQILEDH
NKQLESQLHRLRQLLEQPQERAKVNGTTVSSPSTSLQRDSSQPMLLRVVGSQSTDSDSM
GEEDLLSPPQDTSTGLEEVMEQLNNSFPSSRGRNTPGKPMREDTM"
```

```
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 13977)
AUTHORS     Jones,K.A., Volkmuth,W. and Walker,M.G.
TITLE       Bone remodeling genes
JOURNAL     Patent: US 6426186-A 60 30-JUL-2002;
FEATURES            Location/Qualifiers
source      1..13977
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN

Query Match         99.2%;   Score 1290;  DB 6;  Length 13977;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches 1301;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

Qy      1  CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGGGAATTGAA     60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8716  CGACTTCCAGCAGTTCAGAAGCAGAACGATGTACATAGGGCCTTCAGAGGGAATTGAA   8775

Qy     61  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACGTACGACGTGACGAGAGCA    120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8776  AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACGTACGACGTGACGAGAGCA   8835

Qy    121  GCCTTTGAAGGACTAGAGAAACTCTCACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG    180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8836  GCCTTTGAAGGACTAGAGAAACTCTCACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAG   8895

Qy    181  AGCCCAGAAATGTCACTCGGCTTCACGAAAGCAGGTGAGGAGGTCAATACTGAGTGGGA    240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8896  AGCCCAGAAATGTCACTCGGCTTCACGAAAGCAGGTGAGGAGGTCAATACTGAGTGGGA   8955

Qy    241  AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT    300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8956  AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACT   9015

Qy    301  CCAGGAACTTCAAGAGAGCCACGGATGAGCTGGACCTCAAGCTGCCAAGCTGAGGTGAT    360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9016  CCAGGAACTTCAAGAGAGCCACGGATGAGCTGGACCTCAAGCTGCCAAGCTGAGGTGAT   9075

Qy    361  CAAGGGATCCTGGCAGCCCGTGGGCCATCCTCCTCCAAGATCACCTCGA    420
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   9076  CAAGGGATCCTGGCAGCCCGTGGGCCATCCTCCTCCAAGATCACCTCGA   9135

Qy    421  GAAAGTCAAGGCACTTCGAGGAGAAATTGCGGCCTCTGAAAGAGAACGTGAGCCACGTCAA    480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9136  GAAAGTCAAGGCACTTCGAGGAGAAATTGCGGCCTCTGAAAGAGAACGTGAGCCACGTCAA   9195

Qy    481  TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTTCAGCTGCTATAAACCTCAGCAC    540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9196  TGACCTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTTCAGCTGCTATAAACCTCAGCAC   9255

Qy    541  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTGCAGGACCCAGTCAG    600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9256  TCTGGAAGACCTGAACACCAGATGGAAGCTTCTGCAGGTGCCGTGCAGGACCCAGTCAG   9315

Qy    601  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCACTTTCTTTCCACGTC    660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9316  GCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCACTTTCTTTCCACGTC   9375

Qy    661  TGTCCAGGGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA    720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9376  TGTCCAGGGGTCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA   9435

Qy    721  CGAGACTCAAACAACTTGCTGGGACACCCACATCCCAAAAATGACAGAGCTCTACCAGTCTTTAGC    780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9436  CGAGACTCAAACAACTTGCTGGGACACCCACATCCCAAAAATGACAGAGCTCTACCAGTCTTTAGC   9495

Qy    781  TGACCTGAATAATGTCAGATTCTCAGCTTATAGAGACTGCCATGAAACCTCCGAGACTGCA    840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9496  TGACCTGAATAATGTCAGATTCTCAGCTTATAGAGACTGCCATGAAACCTCCGAGACTGCA   9555

Qy    841  GAAGGCCTTTGCTTGGATCCTCTGAGCCTGTCAGCTGCGCATGTGATGCCTTGACCAGCA    900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9556  GAAGGCCTTTGCTTGGATCCTCTGAGCCTGTCAGCTGCGCATGTGATGCCTTGACCAGCA   9615

Qy    901  CAACCTCAAGCAAAATGACCAGCCCATGGATATCTGCAGATATCTGCAGATTATTAATTGTTTGACCAC    960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9616  CAACCTCAAGCAAAATGACCAGCCCATGGATATCTGCAGATATCTGCAGATTATTAATTGTTTGACCAC   9675

Qy    961  TATTTATGACCGCCTGGAGCAAGAGCAACAACAATTGGTCAAGTCCCTCCTCTGCGTGA   1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9676  TATTTATGACCGCCTGGAGCAAGAGCAACAACAATTGGTCAAGTCCCTCCTCTGCGTGA   9735

Qy   1021  TATGTGTCTGAACTGGCTGCTGCGTGAATGTTTATGATCACGGACGACAGGAGGATCCGTGT   1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9736  TATGTGTCTGAACTGGCTGCTGCGTGAATGTTTATGATCACGGACGACAGGAGGATCCGTGT   9795

Qy   1081  CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG   1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9796  CCTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGACAAGTACAG   9855

Qy   1141  ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCCGCAGGCT-GGGCC   1199
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   9856  ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCCGCAGGCTGGGGCC   9915

Qy   1200  TCCTTCTGCATGATTCTATCCAAATTCCAAGACACAGTTGGGTGAAGTTGCACTCTTTGGG   1259
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9916  TCCTTCTGCATGATTCTATCCAAATTCCAAGACACAGTTGGGTGAAGTTGCACTCTTTGGG   9975

Qy   1260  GCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG   1301
           ||||||||||||||||||||||||||||||||||||||||||
Db   9976  GCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTG   10017
```

```
RESULT 13
AF070485
LOCUS       AF070485                13887 bp    mRNA    linear    MAM 09-DEC-1998
DEFINITION  Canis familiaris dystrophin mRNA, complete cds.
ACCESSION   AF070485
VERSION     AF070485.1   GI:3982750
KEYWORDS    .
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1 (bases 1 to 13887)
  AUTHORS   Carville,K.S., Mann,C.J., Schatzberg,S.J. and Wilton,S.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUN-1998) ANRI, Pathology, University of Western
            Australia, Verdun Street, Nedlands, WA 6018, Australia
FEATURES            Location/Qualifiers
     source          1..13887
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /strain="breed Golden retriever"
                     /db_xref="taxon:9615"
     5'UTR           <1..289
     CDS             290..11132
                     /codon_start=1
                     /product="dystrophin"
                     /protein_id="AAC83646.1"
                     /db_xref="GI:3982751"
                     /translation="MLWWEEVEBDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDL
                     QDGRRLLDLLEGLTGQKLPKEKGSTRVIALNNVNKALRVLQKNNVDLVNIGSTDIVDG
                     NHKLTLGLIWNIILHWQVKNVMKNIMAGLQQTNSEKILLSWVRQSTRNYPQVNVINFT
                     TSWSDGLALNALIHSHRPDLFDWNSVVCQQSATQRLEHAFNIAKYQLGIEKLLDPEDV
                     ATTYPDKKSILMYITSLFQVLPQVSIEAIQEVEMLPRPSQVTREBHFQIHQMHYSQ
                     QITVSLAQGYERAPSFPKPRFKSYAYIQAAVVTTSDDFTRSPLPSQHLETPEDKSFGRS
                     LTETEANLDSYQTALEEVLSWLLSAEDALQAQGEISNDVVEBHQPHTHEGYMMDLTS
                     HQGRVGNVILQLGSOLIGTGKLSEDEETBVQEQMNLLNSRWECLRVASMEKQSNLHKVL
                     MDLQNQQLKELNDWLTKTEERTRKMEEPLGPDIEDLKRQVQQHKVLQEDLEQEQVRV
                     NSLTHMVVVDESSGDHATAALEEQLKVLGDRWANICRWTEDRWVLLQDILLKWQRFT
                     EEQCLFSAWLSEKEDAVNKIHTGFKDQSEVLSNLQLAVLKTDLEKKKQTMDKLCSL
                     NQDLLSALKNTVVAHKMEAWLNFAQRWDNLVQKLEKSSAQISQAVTTQPSLITQTTV
```

This page contains a protein/DNA sequence alignment printout (rotated) consisting of a translated query protein sequence, paired Qy/Db nucleotide alignment blocks numbered by position, and the following legible structural annotations:

3'UTR   11333 ..>13887
repeat_region 11789 ..11828
              /rpt_type=tandem
              /rpt_unit="ca"

ORIGIN

Query Match   88.7%; Score 1154.4; DB 4; Length 13887;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1209; Conservative 0; Mismatches 0; Indels 91; Gaps 0;

Qy   1262 AGTAACATTGAGCCAGTGTCCGGAGCTGCTTCCAATTTG 1301
           ||||||||||||||||||||||||||||||||||||||
Db 10043 AGTAACATTGAGCCGAGTGTCAGGAGCTGCTTCCAGTTTG 10082

RESULT 14
AX306153
LOCUS   AX306153   13815 bp   DNA   linear   PAT 11-DEC-2001
DEFINITION Sequence 904 from Patent WO0188188.
ACCESSION   AX306153
VERSION   AX306153.1 GI:17645441
KEYWORDS
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
  TITLE   Method for examining ischemic conditions
  JOURNAL   Patent: WO 0188188-A 904 22-NOV-2001;
            School Juridical Person Nihon University (JP)
FEATURES   Location/Qualifiers
     source   1..13815
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

Query Match   87.3%; Score 1135.2; DB 6; Length 13815;
Best Local Similarity   92.1%; Pred. No. 0;
Matches 1197; Conservative   0; Mismatches 103; Indels   0; Gaps   0;

Qy   2 GACTTTCCAGCAGTTCAGAAGCAGAACGATGTACATTAGGGCCTTCAAGAGGGAATTGAAA 61
           |||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||||||
Db   8708 GATTTCCAGCAGTTCAGAAGCAGAATGATTATACATAGGGCCTTCAAGAGGGAATTGAAA 8767

Qy   62 ACTAAAGAACCTGTAATCATGAGTACTCTTGGAGACTGTGAGAGATATTTCTGACAGAGCAG 121
           ||||||||||||||||||||||||| ||||||||||||||| ||||||||||||||||||||
Db   8768 ACTAAAGAACCTGTAATCATGAGTACTCTGGGAGACTGTGAGAGATATTTCTGACAGAGCAG 8827

Qy   122 CCTTTGGAAGGACTAGAGAAAACTTACCAGGAGGACCCAGAGAACTGCCTCCTGAGGAGAGA 181
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   8828 CCTTTGGAAGGACTAGAGAAAACTCTACCAGGAGGACCCAGAGAACTGCCTCCTGAGGAGAGA 8887

Qy   182 GCCCAGAATGTCACTCGGCTTCTACGAAAGCAGCTGAGGAGGTCAATACTGAGTGGGAA 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8888 GCTCAGAATGTCACTCGGCTCCTACGAAAGCAGCTGAGGAGGTCAAGAGGTCAATGGGAC 8947

Qy   242 AAATTGAACCTGCGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCTGAAAGACTC 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8948 AAATTGAACCTGCGCAGCTGATTGGCAGAGAAAAATAGATGAAGATGAAGCTCTTGAAAGACTC 9007

Qy   302 CAGAACTTCAAGAGGCCACGTGAGCGGATGAGCTGGACCTCAAGCTGAGGTGATC 361
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   9008 CAGAACTTCAAGAGAAGCTTCAGGGGAGAAATTGCACCTCTTAAAGAGAAGTCAATCGTGTCAAT 9067

Qy   362 AAGGGATCCTGGCAGCCCGTGGGCGATCTCTCCATTGACTCCTCCGAGATCACCTCGAG 421
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9068 AAGGGATCCTGGCGAGCCCGTGGGGGATCTCCTCATTGACTCCTCTGCGAAGATCACCTTGAA 9127

Qy   422 AAAGTCAAGGCACTCGAAGGAGGAGAAATTGCGCCTCGAAAGAAGAAGCTGAGCCACGTCAAT 481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9128 AAAGTCAAGGCACTTCGGGGAGAAATTGCACCTCTTAAAGAGAAGTCAATCGTGTCAAT 9187

Qy   482 GACCTTGCTTCGCCAGCTTACCACTTTGGGCATTCAGCTCTTCAGCTCCTATAACCTCAGCACT 541
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9188 GACCTTGCACATCAGCGTCGACCACACCAGATGGAAGTTCTGCAGTCAGGTCTCCAGGAGGACC... 9247

Qy   542 CTGGAAGACCTGAACACCAGATGGAAGTTCTGCAGTGGCCGTCGAGGACCGAGTCAGG 601
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9248 TTGGAAGATCTGAATACCAGATGGAAGTTCTACAGTGAGGCCATCCAGCACTTCTTCCACGTCT 9307

Qy   602 CAGCTGCATGAAGCCACCAGGGACTTTGGTCCAGCATCCAGCACTTTCTTCCACGTCT 661

Db   9308 CAGCTGCATGAAGCCCCACCAGGGACTTTGGTCTGCATCCCAGCACTTCCTTCCACTTCA 9367
           ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Qy   662 GTTCAGGGTGCCCTGGGAGAGAGCCCATCCGCCAAACAAAGTGCCTACTATATCAACCAC 721

Db   9368 GTTCAGGGTGCCCTGGGAGAGAGCCCATCTGCCAAACAAAGTGCCTACTATATCAACCAC 9427
           ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy   722 GAGACTTCAAACAACTTGCTGGACACATCCCAAAATGACACAGAGCTTCTACCAGTCTTTAGCT 781

Db   9428 GAGACCCAAACCACTTGTTGGACACCACCCAAAATGACACAGAGCTTCTACCAGTCTTTTAGCT 9487
           ||||| ||||| ||||| ||||||| ||||||||||||||||||||||||||||||||| ||||
Qy   782 GACCTGAATAATGTCAGATTCAGCTTATAGGACTGCAGCTGCATGTGCATGTGATGCCTGGACCAGCAG 841

Db   9488 GACCTGAATAATGTCAGGTTCAGGTTCTCCCGTATAGGACTGTCAGCTGTCAGTGCATGTGATGCCGGCAGCAC 9547
           ||||||||||||||||| ||||| |||| ||||||||||||| |||||||||||||||||||| |||| ||
Qy   842 AAGGCCCTTGCTTGGGATCTCTTTGAGCCTTGGTCTGCTGGGATGCTGGCTCGAAGACTGCAG 901

Db   9548 AAGGCCCTTGCTTGGGATCTCTTTGAGCCTTGGTCTGCTGGCCTGGACCAGCAGCAC 9607
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Qy   902 AACCTCAAGCAAAATGACCAGCCCATGGAATCTCGACAGATTATTAATTGTTTGCACCACT 961

Db   9608 AACCTCAAGCAAAATGACCAGCCCATGGAATCTCCGCGATAATCCTGCAGATAATTAACTGTTTGACTACA 9667
           ||||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Qy   962 ATTTATGACCGGCTGGAGCAAGAGACACAACAATTTGGTCAACGTCCCTCTCGGCGTGGAT 1021

Db   9668 ATTTATGATCGTCTGGAGCAAGAGCACAAGAGCACAACAATCTGGTCAATGTCCCTCTCGTGTGGAT 9727
           ||||||||| || |||||||||||| ||||| ||||||| |||||| |||||||||||||||||
Qy 1022 ATGTGCTTGAACTGGCTGCTGAATGTTTATGATCACGGCGACGAACGGGAGGATCCGTGTC 1081

Db   9728 ATGTGTCCTCAACTGGCTTCTCAATGTTTATGATCACGGCGACGAACAGGGAGGAGGATCCGTGTC 9787
           ||||| ||||||||||| |||||||||||||||||||||||||||| |||||||||||||||||
Qy 1082 CTGTCTTTAAAACTGGCATCATTCCCTGTGTAAAGCACATTGGAAGCACAAGTACAGA 1141

Db   9788 CTGTCTTTAAAACTGGCATCATTCTCTGTGTAAAGCACACTGGAAGCACAAGTACAGA 9847
           |||||||||||||||||||||||| |||||||||||||| |||||||||||||||||
Qy 1142 TACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAGCGCAGCTGGCCTC 1201

Db   9848 TACCTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTGTGTGACCAGCGTAGGCTGGTCTT 9907
           |||||||||||||||||||||||||||| |||| | ||||||||| ||||| |||| ||
Qy 1202 CTTCTGCATGATTCTATCCAAATTCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGGGC 1261

Db   9908 CTTCTGCATGATTCATTCATTCAAATCCCAAGACAGTTGGGTGAAGTTGCTTCCTTTGGGGGC 9967
           ||||||||||||| || | |||||| | |||||||||||||||||||||| |||||||||||
Qy 1262 AGTAACATTGAGCCAGTGTCGCCGAGTGTCAGGAGCTGCTTCCAATTTG 1301

Db   9968 AGTAACATTGAGCCGAGTGTCAGGAGCTGCTTCCAATTTG 10007
           |||||||||||||| ||||||||||||||||||||||||

RESULT 15
AX538582
LOCUS   AX538582   13815 bp   DNA   linear   PAT 23-NOV-2002
DEFINITION Sequence 2 from Patent WO0229056.
ACCESSION   AX538582
VERSION   AX538582.1 GI:25271088
KEYWORDS
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Chamberlain,J.S. and Harper,S.Q.
  TITLE   Mini-dystrophin nucleic acid and peptide sequences
  JOURNAL   Patent: WO 0229056-A 2 11-APR-2002;
            THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES   Location/Qualifiers
     source   1..13815
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

Query Match   87.3%; Score 1135.2; DB 6; Length 13815;
Best Local Similarity   92.1%; Pred. No. 0;

Matches 1197; Conservative   0; Mismatches 103; Indels   0; Gaps   0;

```
Qy      2 GACTTTCCAGCAGTTCAGAAGCAGGACGATGTACATAGGGCCTTCAAGAGGGAATTGAAA 61
          ||||||||||||||||||||||||||||||| |||||| || ||||||||||||||||||
Db   8708 GATTTCCCAGCAGTTCAGAGAGCAGAAGATGATATACATAGGGCCTTCAAGAGGGAATTGAAA 8767

Qy     62 ACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACAGAATATTCTGACAGAGCAG 121
          |||||||||||||||| ||||||||||||||||||||||| ||||| ||||||||||||||
Db   8768 ACTAAAGAACCTGTAATCATGAGTACTCTGGAGACTGTGAGAATATTTCTGACAGAGCAG 8827

Qy    122 CCTTTGGAAGACTAGAGAAACTCTACCAGGAGAGCTGCCTCCTGAGGAGAGA 181
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   8828 CCTTTGGAAGGACTAGAGAAACTCTACCAGGAGAGCCCAGAGAACTGCCCTCCTGAAGAAAGA 8887

Qy    182 GCCCAGAATGTCACTCGGCTTCACGGAAGCAGGTCAATACTGAGTGGGAA 241
          ||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   8888 GCTCAGAATGTCACTCGGCTCCTACGAAACGCAGGTCAAACGCTGAGTGGGAC 8947

Qy    242 AAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCTTGAAAGACTC 301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8948 AAATTGAACCTGCGCTCAGCTGATTGGCAGAGAAAATAGAGCTTTGAAAGACTC 9007

Qy    302 CAGGAACTTCAAGAGGCCACGGATGAGCTGACCTCAAGCTGCCCAAGCTGGGTGATC 361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9008 CAGGAACTTCAAGGAAGCTGCCATGAGCTGACCTCAAGCTGCCCAAGCTGAGGTGATC 9067

Qy    362 AAGGGATCCTGCAGCCGTGGCCGATCTCCTCCTCCAAGATCACCTCGAG 421
          |||||||||||||||||||||||||||||||||||||||||||||| ||
Db   9068 AAGGGGATCCTGCAGCCAGTGGGGGATCTCCTCTCCTGCAAGATCACCTTGAA 9127

Qy    422 AAAGTCAAGGCACTTCGAGAGGAAATTGCGCCCTCTGAAAGAGAACGTGAGCCACGTCAAT 481
          ||||||||||||||||||||||||||||| |||||||||||||||| ||||||| |||||
Db   9128 AAAGTCAAGGCACTTCGGGGAGAAATTGCACCTCTAAAGAGAATGTCAATCGTGTCAAT 9187

Qy    482 GACCTTGCTCGCCAGTCTACCACTTGGTCCAGCATTCAGCCTCTCACCGTATAACCTTCAGCACT 541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9188 GACCTTGCAATCAGCTGACCACTGGCATTCAGCCTCTCACCTTATAACCTTCAGCACT 9247

Qy    542 CTGGAAGACCTGAACACCAGATGGAAGCTTCTGCGAGGTGGCCGTCGAGGACCGAGTCAGG 601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9248 TTGGAAGATCTGAATACCAGATGGAAGCTTCTACAGGTGGTGTGGAGGACCGTGTCAGA 9307

Qy    602 CAGCTGCATGAGAGCCCACAGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCT 661
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9308 CAGCTGCATGAGAGCCCACAGACTTTGGTCCTGCCATCCGCATCCCAGCCACTTCCTTTCCACTTCA 9367

Qy    662 GTCCAGGGTCCCTGGGAGAGAGCCATCTCGGCCAAACAAAGTGCCCATACTATATCAACCAC 721
          |||||||||||||||||||||||||||||| ||||| |||||||||||| ||| |||||||
Db   9368 GTTCAGGGTCCCTGGGAGAGAGCCATCTCCTGCCAAACAAAGTGCCCTACTATATCAACCAC 9427

Qy    722 GAGACTTCAAACAACTTGCTGGGACCATCCCAAAATGACACAGAGCTCTACCAGTCTTTAGCT 781
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||| |||||
Db   9428 GAGACCCAAACACACTTGTTGGGACCACCCCAAAAATGACAGAGCTCTACCAGTCTTTAGT 9487

Qy    782 GACCTGAATAATGTCAGATTCTCAGCTTCCCGGTTATAGGACTCGCCATGGACTGCGAAACTCCGAAGACTGCAG 841
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9488 GACCTGAATAATGTCAGGTTCTCCCGTATAGGACTGCCATGGACTGCGAAGCTCAGAAGGCTCCAG 9547

Qy    842 AAGGCCCTTTGCTTGGATCTCTTGGATCTGAGCCTGCATGTGATGCCTTGGACCAGCAC 901
          ||||||||| |||||||| ||||| ||||||||||||||||||||||||||||||||||
Db   9548 AAGGGCCCTTTGCTTGGATCTCTTGGATCTGAGCCTGCATGTGATGCCTGGACCAGCAC 9607

Qy    902 AACCTCAAGCAAAATGACACCAGCCATGAATCCTGCAGATTATTAATTGTTTGACCACT 961
          ||||||||||||||||||||||||||||||||||| |||||| |||||||||||||||
Db   9608 AACCTCAAGCAAAATGACCACCGCCATGAATCCTGCAGATAATTTAACTGTTTGACTACA 9667

Qy    962 ATTTATGACCGCCTGAGCAAGAGCACACAACAATTGGTCAACGTCCCTCTCTGGGTGAT 1021
          ||||||||||||||||||||||||||||||| ||||| |||||||||||||||||||||
Db   9668 ATTTATGATCGTCTGGAGCAAGAGCACAACAATCTGGTCAATGTCCCTCTCTGTGTGGAT 9727

Qy   1022 ATGTGTTCTGAACTGGCTGCTTCAATGTTTATGATACCGGTGTC 1081
          ||||||||||||||||||||||||||||||||||||||||||||
Db   9728 ATGTGTTCTCAACTGGCTTCAATTGTTTATGATACCGGTGTC 9787

Qy   1082 CTGTCTTTTAAAACTGGCATCATTTCCCGTGTAAAGCACATTTGGAAGACAAGTACAGA 1141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9788 CTGTCTTTTAAAACTGGCATCATTTCTCTGTGTAAAGCACACTTGGAAGCACAAGTACAGA 9847

Qy   1142 TACCTTTCAAGCAAGTGGCAAGTTCAAACAGGATTTTGTGACCAGGATTTTGTGACCAGCGCCAGGCTGGGCCTC 1201
          ||||| ||||||||||||||| ||||||||| | |||||||||||||||||||| | ||
Db   9848 TACCTTTCAAGCAAGTGGCAAGTTCAACTGGCTTTTGTGACCAGCGTAGGCTGGGCTGGTCTT 9907

Qy   1202 CTTCTGCATGATTCTATCCAAATTCCAAGACAGTGGGTGAAGTTGCATCCTTTGGGGGC 1261
          |||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||
Db   9908 CTTCTGCATGATTCATTCATTCAAATCCCAAGACAGTGGGTGAAGTTGCTTCCTTTGGGGGGC 9967

Qy   1262 AGTAACATTGAGCCAAGTGTCCGAGCTGCTTCCAATTTG 1301
          |||||||||||||||||||||||||||||||||||||||
Db   9968 AGTAACATTGAGCCGAGTGTCAGGAGCTGCTGCCAATTTG 10007
```

Search completed: April 5, 2004, 02:39:55
Job time : 3394.8 secs